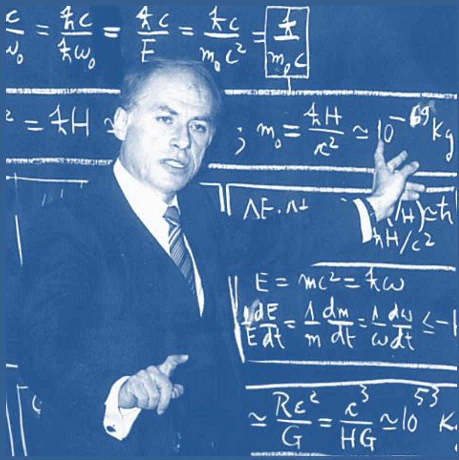
Ioan-Ioviț Popescu

Optics

I. Geometrical Optics



Edited by
Lidia Vianu

Translated into English by
Bogdan Voiculescu

Contemporary Literature Press

http://editura.mttlc.ro
The University of Bucharest. 2017

Wednesday 31 May 2017 Online Publication

Press Release

Ioan-Ioviț Popescu

Optics.
I. Geometrical Optics

Translated into English by **Bogdan Voiculescu**

ISBN 978-606-760-105-3

Edited by **Lidia Vianu**

Ioan-Ioviț Popescu's *Geometrical Optics* is essentially a book about light rays: about what can be seen in our universe. The name of the author is well-known to physicists all over the world. He is the scientist who **predicted the Etheron in the year 1982**. We have already published the book of that unique discovery: *Ether and Etherons. A Possible Reappraisal of the Concept of Ether*,

http://editura.mttlc.ro/iovitz-etherons.html .

What students have to know about the book we are publishing now is that it was

Optica geometrică de Ioan-Ioviț Popescu se ocupă de buna aproximație a luminii sub forma de raze (traiectorii). Numele autorului le este cunoscut fizicienilor din lumea întreagă. Ioviț este acela care a prezis existența **Etheronului** încă din anul **1982**. Editura noastră a publicat nu demult cartea acelei descoperiri absolut unice: *Ether and Etherons. A Possible Reappraisal of the Concept of Ether*,

http://editura.mttlc.ro/iovitz-etherons.html .

Studenții facultății de fizică din Bucureşti ştiau în anul 1988—aşa cum ştiam şi eu de la autorul însuşi—că această carte a fost îndelung scrisă de mână de către

handwritten by its author over and over again, dozens of times—formulas, drawings and everything. It was meant to be published as such, in facsimile, and its cover was intentionally black. It was so much used by students at the library of the Romanian Faculty of Physics, that all the copies are almost illegible now...

Very soon, we also mean to publish the original book as facsimile in Romanian. For the time being, we are offering students all over the world its English version, reminding them that Ioan-Ioviț Popescu is the first physicist who brought to light and gave a name to the smallest particle in our universe: the Etheron.

autor cu scopul de a fi publicată în facsimil: ea a fost rescrisă de la capăt la fiecare nouă corectură, şi au fost refăcute de nenumărate ori figurile, forumulele, tot ce era pe pagină. Cartea a fost atât de mult folosită încât exemplarele de la biblioteca facultății abia dacă se mai pot desluşi... Pentru noi, cei izolați în comunism, această carte sfida, odată cu tiparnița, un întreg *1984* devenit realitate. Coperta ei era, în mod semnificativ, neagră.

Urmează să publicăm în curând şi volumul olograf în limba română. Oferim acum studenților din întreaga lume versiunea lui în limba engleză. Le reamintim cu această ocazie că, primul din lume, Ioan-Ioviț Popescu a argumentat şi a calculat cea mai mică particulă din univers, numind-o Etheron.

Lidia Vianu

Ioan-Ioviț Popescu

Optics

I. Geometrical Optics



Edited by
Lidia Vianu

Translated into English by
Bogdan Voiculescu

CONTEMPORARY
LITERATURE PRESS

http://editura.mttlc.ro
The University of Bucharest. 2017

CONTEMPORARY LITERATURE PRESS
The online Publishing House of the University of Bucharest
http://editura.mttlc.ro
Editura pentru studiul limbii engleze prin literatură

Lidia Vianu
Director
C. George Sandulescu
Executive Advisor

UNIVERSITATEA DIN BUCUREŞTI · BRITISH COUNCIL · Ambasáid na hÉireann / Embassy of Ireland / Ambasada Irlandei · INSTITUTUL CULTURAL ROMÂN

ISBN 978-606-760-105-3

Editing, Cover Design and overall Layout by
Lidia Vianu

Advisor for the English version: Ioan-Ioviţ Popescu
Proofreading: Ioan-Ioviţ Popescu, Cristian Vîjea
IT Expertise: Cristian Vîjea
PR Manager: Violeta Baroană

Acknowledgements

This book was first published in Romanian, as Ioan-Ioviţ Popescu: *Optica. I. Optica geometrică*, in the year 1988, at Universitatea din Bucureşti, Facultatea de Fizică. It was a facsimile of the author's handwritten manuscript, which had been meant from the very beginning to appear in the author's handwriting. The facsimile edition in Romanian will soon be published by CLP.

The second edition of this book appeared in 2006, as *Optica geometrică*, by Ioan-Ioviţ Popescu and Florea Uliu, published by Editura Universitaria. This second edition included Professor Uliu's "Alte aspecte fizice şi fiziologice privind funcţionarea ochiului uman ca receptor optic" [pp 57-64]. Professor Uliu also scanned there all the geometrical figures that had previously been published in facsimile in Ioan-Ioviţ Popescu's book, and which he has had the kindness to forward to us for the present edition.

Ioan-Ioviț Popescu

Optics

I. Geometrical Optics

Edited by
Lidia Vianu

Translated into English by
Bogdan Voiculescu

CONTEMPORARY LITERATURE PRESS

http://editura.mttlc.ro

The University of Bucharest. 2017



Table of Contents

Contemporary
Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017



PREFACE

The present volume comprises the first part of the two-semester Optics Course intended for use by the students of the Faculty of Physics of Bucharest University. In this volume we discuss geometrical optics, while the subsequent volumes are dedicated to wave optics and quantum optics.

Optics is one of the earliest sciences, having originated in antiquity, and many of the current concepts and representations of physics are linked to the effort towards understanding the essence of light-related phenomena. Although nowadays geometrical optics appears to be a limit case of Maxwell's equations applied to the phenomena of propagation of electromagnetic fields of very short wavelenghts, the design of optical instruments is mainly based on the behaviour of light rays considered through these systems, since it rarely happens that diffraction bypasses geometrical aberrations. Also, the study of finer wave phenomena, such as interference, polarization, and diffraction, always requires preliminary observation of the geometrical path of light rays.

In Chapter I we introduce the principles of geometrical optics in considering the eikonal equation for wavefronts and the associated equation for light rays. Subsequently, we again take up these issues from a global perspective, and elaborate Fermat's principle and Lagrangian formalism for light rays. These general laws will then allow us to construct the geometrical theory of optical images, based on the concept of stigmatism.

In Chapter II we deal with the important class of rotationally symmetrical dioptric systems. In order to study the geometrical behaviour of paraxial rays, we introduce the transfer-matrix method, which we use in a detailed description of the main optical instruments. Subsequently we give an account of the aspects which limit the effectiveness of optical apparatuses, and discuss the methods of adjusting for chromatic and geometrical aberrations.

In Chapter III we give a study of phenomena and applications where the refractive index varies continuously. We describe the trajectory of rays in systems with planar, cylindrical and spherical symmetry, with application in integrated optical circuits and fiber optics.

We complete the present study of the methods of geometrical optics with an account of the relevant moments in the evolution of concepts pertaining to the field of optics. We give special attention to Snell and Descartes's establishing the law of refraction, to Fermat enunciating the principle of least time, we discuss the historical

context in which the telescope and microscope were invented, thus opening up the world of very long and very short distances. We conclude the historical overview in this volume with Fraunhofer making spectral observations for the first time, which marked the beginning of the way to elucidating the atomic structure and the nature of light.

Măgurele,
15th of March 1988



Chapter I
THE PRINCIPLES OF GEOMETRICAL OPTICS

It is a well-known fact that, through his sense organs, man is permanently connected to the environment in which he lives and works. It is with their help that man obtains the entirety of his information regarding surrounding objects and phenomena. It has been estimated that about 90% of the information a human being receives and processes throughout his lifetime is gained visually. The transfer of visual information from "objects" (located close by, as is the case most times, or far away, as in the case of astronomical observations) to their "observers" takes place at ultrahigh speed, through certain radiations known as ***luminous radiations***, or simply ***light***.

Although throughout the ages the science of light, that is, ***optics***, has been approached by great personalities of natural science, such as Huygens, Newton, Young, Fresnel, Maxwell, Einstein, Feynman, of philosophy, such as Descartes, Spinoza, and also of the arts, such as Leonardo da Vinci and Goethe, its evolution is not linear. Whereas up to the first decades of the 20th century there was a long, often circuitous period of factual and conceptual accumulations, the last few decades has seen optics becoming one of the most dynamic parts of physical science.

In this book concerning the fundamentals of optics, we approach the issues of ***geometrical optics***, which is based on the ***simplest model of light propagation***. As it is known, geometrical optics is that part of optics in which the propagation of light and its interaction with material media are studied using the concept of ***light ray***, defined as a curve (and, particularly, a straight line) along which luminous energy is propagated. This concept appeared and was founded on a phenomenological basis. It began with the observation of shadows and penumbrae, as well as of the images formed in the camera obscura.

Light beams are considered to be made up of an infinite set of ***independent*** light rays, each ray being propagated in a straight line in homogenous media, and behaving according to the well-known laws of reflection and refraction at the point of separation between two different media.

Because of the important place modern optical technology occupies in designing, as well as in producing various types of components, instruments or machines, the optics of light beams, that is, geometrical optics, is and will always be a distinct branch of optics, irrespective of the level at which one approaches it.



Today, despite all its limitations, geometrical optics is believed to embody the specific features of a scientific theory, because it has a logical, unitary structure, given by its fundamental principle, ***Fermat's principle***, from which are derived all the laws and consequences subjected to experimental analysis.

Although light propagation may be treated in detail with the help of Maxwell's equations and the corresponding equation of electromagnetic waves, many problems of a practical nature are much more easily solved using the concept of light ray and the laws of geometrical optics. As we will soon show, geometrical optics, or light ray optics, is an approximation of wave optics for λ wavelengths that are very small (theoretically, for $\lambda \to 0$) compared to the dimensions of objects (obstacles) limiting the beams of light. In this approximation, energy propagates along the rays of light, defined as the multitude of trajectories normal to the wavefronts. Wavefronts are defined as the loci of points of equal phase.

Note that any light ray, as the trajectory of a mathematical point in space, is only a geometrical abstraction, and is not physically observable. In fact, if we tried to isolate a single light ray using a diaphragm of variable diameter, we would see that beyond a certain limit, instead of thinning, the beam enlarges and becomes divergent. This deviation from the propagation of energy along the geometrical rays is due to light's wavelike nature, and is caused by the ***diffraction*** of waves. Still, because of the $\lambda \to 0$ restriction, wave diffraction cannot be described within geometrical optics.

We begin this chapter by deducing the fundamental equation of geometrical optics (***the eikonal equation***) for wavefronts and the associated equation for light rays.

1.1 The Eikonal and the Light Ray Equations

Let us consider a transparent and isotropic optical medium and the ***scalar equation of waves***,

(1) $$\Delta E = \frac{1}{v^2} \cdot \frac{\partial^2 E}{\partial t^2},$$

where $E(\vec{r}, t)$ represents any of the scalar components of the electromagnetic field, and $v(\vec{r}) = c/n(\vec{r})$ is the speed of light in the point considered within the medium of refractive index $n(\vec{r})$. Note that equation (1) is valid only if the variations of n and of his gradient along a wavelength are negligible, conditions which are ideally met within the limits of geometrical optics ($\lambda \to 0$).

We will now only consider the propagation of monochromatic waves, that is, of waves temporally dependent according to the factor $exp(-i\omega t)$. In this case, $\partial^2 E/\partial t^2 = -\omega^2 E$, and equation (1) becomes the equation for monochromatic waves (or Helmholtz's equation),

(2) $$\Delta E + k^2 E = 0,$$

where $k = \omega/v = 2\pi/\lambda = k_0 n$ and $k_0 = \omega/c = 2\pi/\lambda_0$ are the moduli of wave *vector*s in a medium and in a vacuum, respectively, and $\lambda = \lambda_0/n$.

The most important solutions to equation (2) ***in homogenous media*** ($n = const.$) are planar, cylindrical, and spherical waves. The more complicated waves can be represented as superpositions of these. Let us first consider monochromatic planar waves in complex representations, that is, solutions of the form

(3) $$E(\vec{r}, t) = E_0 e^{i(\vec{k}\cdot\vec{r} - \omega t)},$$

where E_0 is a constant, usually complex amplitude, $\vec{k} = k\vec{\tau} = k_0 n\vec{\tau}$ is the wave (or propagation) vector, and $\vec{\tau}$ is the unit vector for the direction of wave propagation. Vectors $\vec{\tau}$ and $\vec{k}$ are constant and perpendicular to the wavefronts, which in this case are the planes given at any moment t by the equation $\vec{k}\cdot\vec{r} = \omega t + const.$ Fig. 1 illustrates the position at successive moments of the equiphase plane $\vec{k}\cdot\vec{r} = \omega t$ and the corresponding plane at zero radians $\vec{\tau}\cdot\vec{r} = vt$. The light rays are rectilinear, along the direction $\vec{\tau}$.

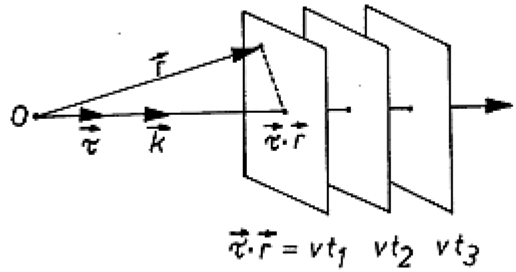


Fig. 1 Three successive positions of the equiphase plane $\vec{k}\cdot\vec{r} - \omega t = 0$, meaning $\vec{\tau}\cdot\vec{r} = vt$.

In non-homogenous media, the refractive index varies spatially, that is, $n = n(\vec{r})$, and expression (3) no longer represents a solution to the wave equation. We will look for harmonic solutions of the more general form:

(4) $$E(\vec{r}, t) = E_0(\vec{r}) e^{i[k_0 \phi(\vec{r}) - \omega t]},$$

where the real scalar function $\phi(\vec{r})$, which represents the spatial phase component, is

called ***eikonal***. The term was introduced by H. Bruns (1895), and comes from the Greek word εἰκών, which means image. The surfaces of constant phase are described

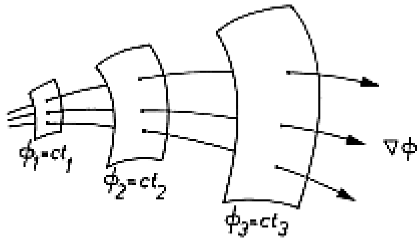


Fig. 2 Three successive positions for the zero radian equiphase surface $\phi(\vec{r}) = ct$ and the associated orthogonal trajectories of light rays in a non-homogenous medium.

at any given moment t by the equation $k_0\phi(\vec{r}) = \omega t + const.$, so that $d\phi = cdt$. Fig. 2 illustrates the position at successive moments of the zero radian equiphase surface $\phi(\vec{r}) = ct$ and the associated orthogonal trajectories $\nabla\phi$ of light rays, which in non-homogenous media are usually curvilinear. As we will soon see, the difference $\phi_2 - \phi_1 = c(t_2 - t_1)$ is sinonymous to the ***optical path*** traversed by the light rays between the considered wavefronts, and is, of course, proportional to the corresponding phase difference $k_0(\phi_2 - \phi_1)$.

Let us determine the equation for the eikonal function $\phi = \phi(\vec{r})$ with the requirement that expression (4) should be a solution to the wave equation. We have

$$\nabla E = (\nabla E_0 + ik_0 E_0 \nabla\phi)e^{i(k_0\phi - \omega t)},$$

$$\Delta E = \{\Delta E_0 + ik_0[2(\nabla E_0)\cdot(\nabla\phi) + E_0\Delta\phi] - k_0^2 E_0(\nabla\phi)^2\}e^{i(k_0\phi - \omega t)},$$

so that, by making the relevant substitutions in wave equation (2), we get:

(5) $$\left[n^2 - (\nabla\phi)^2\right]E_0 + \frac{\Delta E_0}{k_0^2} + \frac{i}{k_0}[2(\nabla E_0)\cdot(\nabla\phi) + E_0\Delta\phi] = 0,$$

or, if we were to write the real and imaginary parts separately,

(5′) $$\left[n^2 - (\nabla\phi)^2\right]E_0 + \frac{\Delta E_0}{k_0^2} = 0,$$

(5″) $$2(\nabla E_0)\cdot(\nabla\phi) + E_0\Delta\phi = 0.$$

Let us first analyze the consequences of equation (5′), which, within the limits

of geometrical optics $\lambda_0 \to 0$, or $k_0 \to \infty$, becomes the non-homogenous first order and second order differential equation

(6) $$(\nabla\phi)^2 = n^2 \qquad ,$$

which allows us to determine the function $\phi(\vec{r})$ if we know the $n(\vec{r})$ distribution of the refractive index and the boundary conditions. This is the ***eikonal equation***, deduced for the first time by A. Sommerfeld and I. Runge (1911), and it represents the ***fundamental equation of geometrical optics***, since the eikonal function $\phi(\vec{r})$ gives a complete description of the optical field with respect to wavefronts.

We can otherwise describe the optical field with reference to the light rays defined as a group of trajectories normal to the wavefronts (the justification for this definition will result from analysis of equation (5″)). If we were to consider the trajectory of light rays in the parameter form $\vec{r} = \vec{r}(s)$, where the independent parameter s is the trajectory arc length (Fig. 3), the unit vector that determines the direction of light rays at any point is

(7) $$\vec{\tau} = \frac{d\vec{r}}{ds} = \frac{\nabla\phi}{|\nabla\phi|} \qquad ,$$

so that the eikonal equation can also take the following equivalent forms:

(8) $$|\nabla\phi| = n \qquad \text{or} \qquad \frac{d\phi}{ds} = n \qquad ,$$

(9) $$\nabla\phi = n\vec{\tau}$$

(10) $$\nabla \times (n\vec{\tau}) = 0 \qquad ,$$

(11) $$\nabla \times \vec{k} = 0 \qquad ,$$

where $\vec{k} \overset{\text{def}}{=} k_0\vec{\tau}$ is the local wave vector.

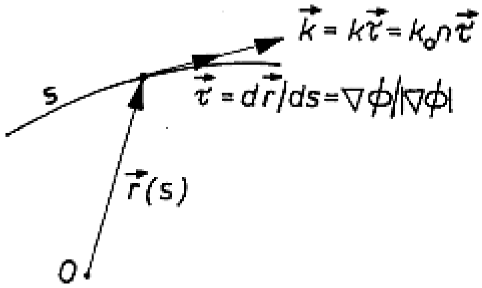


Fig. 3 A light trajectory, in parameter representation $\vec{r} = \vec{r}(s)$.

Note that the graphical integration of the eikonal equation, equation (8), is equivalent to ***Huygens'*** step by step ***construction*** (***Traité de la lumière, 1692***) of wavefronts. Indeed, we have

(12) $$d\phi = nds = cdt \quad ,$$

so $ds = cdt/n(\vec{r}) = v(\vec{r})dt$. In other words, if we consider the points of a wavefront to be synchronous sources of secondary spherical waves, any neighboring wavefront is an envelope for them (Fig. 4). Of course, this construction applies along both directions of propagation (the property of the ***reversibility of light***).

If we differentiate eikonal equation (9) with respect to the s parameter of light trajectory, and take into account equation (8), we get

$$\frac{d}{ds}(n\vec{\tau}) = \frac{d}{ds}(\nabla\phi) = \nabla\left(\frac{d\phi}{ds}\right) = \nabla n \quad ,$$

which is the ***light ray equation***

(13) $$\frac{d}{ds}(n\vec{\tau}) = \nabla n \quad , \text{ or } \quad \frac{d}{ds}\left(n \cdot \frac{d\vec{r}}{ds}\right) = \nabla n \quad .$$



In the special case of homogenous media, we have $n = const.$, $\nabla n = 0$, so that equation (13) becomes $d^2\vec{r}/ds^2 = 0$, and the ray trajectories are the lines given by $\vec{r}(s) = \vec{r}_0 + \vec{\tau}_0 s$, where $\vec{r}_0$ and $\vec{\tau}_0$ are integration constants. Of course, this can also be deduced directly from eikonal equation (10), which becomes $\nabla \times \vec{\tau} = 0$, meaning $\vec{\tau}$ is constant. In the general case of non-homogenous media, the direction of propagation $\vec{\tau}$ changes continually along the ray of light according to the ray equation (13). Since $\vec{\tau}^2 = 1$, we have $\vec{\tau} \cdot (d\vec{\tau}/ds) = 0$, meaning the unit vector $\vec{\tau}$ of direction of propagation and the light ray curvature vector

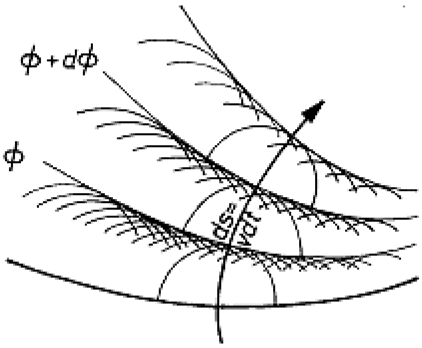


Fig. 4 Huygens' construction

(14) $$\vec{K} = \left(\frac{d\vec{\tau}}{ds}\right) = \frac{\vec{\nu}}{\rho},$$

are orthogonal ($\vec{\nu}$ is the main normal vector unit, and ρ is the local curvature radius), Fig. 5.

The ray equation (13) can also be written as

(15) $$\nabla n = \frac{d}{ds}(n\vec{\tau}) = \frac{dn}{ds}\vec{\tau} + n\frac{d\vec{\tau}}{ds} = \frac{dn}{ds}\vec{\tau} + \frac{n}{\rho}\vec{\nu},$$

a form in which it highlights the coplanarity of vectors $\nabla n, \vec{\tau}, \vec{\nu}$ in the ***osculating plane*** $(\vec{\tau}, \vec{\nu})$. By multiplying scalar equation (15) by $\vec{\nu}$, we get the general expression of the ***light ray curvature***

(16) $$\frac{1}{\rho} = \frac{\vec{\nu} \cdot \nabla n}{n} = \vec{\nu} \cdot \nabla(\ln n),$$

Since it is always true that $1/\rho \geq 0$, we get $\vec{\nu} \cdot \nabla n = |\nabla n| \cos\theta \geq 0$, $\cos\theta \geq 0$, $\theta \leq \frac{\pi}{2}$. Thus, we obtain the ***general rule*** according to which the light ray always curves towards the medium of higher refractive index. Equality corresponds to the limit case of homogenous media ($\nabla n = 0$), where the $1/\rho$ curvature is null, that is, the light ray is rectilinear.

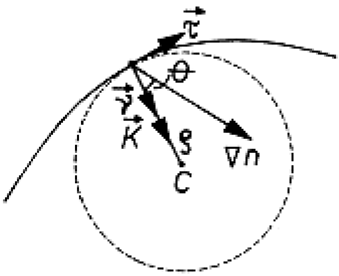


Fig. 5 local curvature of the light ray

Let us analyze equation (5″) in such a case. By referring to equation (9), we will rewrite it as

(17) $$2n\frac{\partial E_0}{\partial s}+E_0\Delta\phi=0$$.

After integration, we get

(18) $$E_0(s)=E_0(0)\exp\left(-\int_0^s\frac{\Delta\phi}{2n}ds\right)$$,

from which it may be deduced that the field amplitude $E_0(s)$ in any point of a given ray depends on an initial value $E_0(0)$ on the ***same*** ray, on the distribution of the refractive index $n(s)$ along the ray, and on the laplacian of the optical path $\phi(s)-\phi(0)=\int_0^s n(s)ds$ (see Chapter 1.2). The equations of geometrical optics do not intercondition the field values on different rays, no matter how close together they may be, so that a light beam appears as an aggregate of independent rays (***the principle of independent light rays***).

Referring again to the eikonal equation $\nabla\phi=n\vec{\tau}$, that is, $\nabla\phi=\nabla\cdot(\nabla\phi)=\nabla\cdot(n\vec{\tau})$, equation (5″) can be rewritten as

(19) $$2(\nabla E_0)\cdot(n\vec{\tau})+E_0\nabla\cdot(n\vec{\tau})=0$$,

or

(20) $$\nabla\cdot\left(E_0^2 n\vec{\tau}\right)=0$$,

that is, as the continuity equation $\nabla\cdot\vec{j}=0$ for stationary incompressible fluids of current density $\vec{j}\sim E_0^2 n\vec{\tau}\sim E_0^2\vec{k}$. Vector $\vec{j}$ is analogous to the Poynting $\vec{S}$ vector of electromagnetic theory, and represents the density of the energy current in the optical field considered in geometrical optics. Out of these considerations results a fundamental concept of geometrical optics, namely that according to which luminous energy propagates along light rays through the tubes of current lines $\vec{j}$.

Using the notation σ for the transverse section area of a narrow beam of light (narrow tube of current lines), as it is used regarding fluids, based on equation (20) we deduce that $E_0^2 n\sigma=const.$ along the beam (tube).

After all these strictly theoretical considerations, we will illustrate how we can ***effectively*** obtain a beam of isolated light rays. Of course, the process entails the introduction of a diaphragm before a spatially extended light wave, of a form expressed in (4), with an arbitrary $\vec{r}$. In order that the rays in the beam should meet the validity conditions of the eikonal equation (small λ_0 wavelength and an amplitude

$E_0(\vec{r})$ with sufficiently slow variation in space, so that it satisfies the inequality $|\Delta E_0/E_0 k_0^2| = (\lambda_0^2/4\pi^2)\cdot|\Delta E_0/E_0| \ll n^2$; for this purpose it is sufficient that $|\lambda_0 \partial E_0/\partial x| \ll E_0$, etc.) this diaphragm must not be too narrow, and the resulting beam must not be too long. Indeed, along the edges of the diaphragm and on the lateral surface of the beam obtained with said diaphragm, the amplitude $E_0(\vec{r})$ varies significantly, and, because of this, the conditions specified above (within parentheses) are not met; as a consequence, the light diffracts, and so considerably widens the beam. The effects of diffraction are insignificant if the diaphragm is wide and if the light beam is short. In diffraction theory, it is demonstrated that the eikonal equation can still be used for distance $s \ll D^2/\lambda_0 \equiv s_0$, where D is the smallest width of the diaphragm. For example, for $\lambda_0 = 500\ nm$ and a circular diaphragm of diameter $D = 1mm$, we get $s_0 = 2m$; this means that, beyond the diaphragm, across a distance of several centimeters, the beam is still composed of independent light rays, to which equations (5″) and (9) still apply. When $D \to \infty$, or $\lambda_0 \to 0$, we have $s_0 \to \infty$, and the approximation of geometrical optics applies to any distances s, however large. Unfortunately, this kind of situations are very rare in practice. That is why geometrical optics is only a "first approximation" of optics.

1.2 Fermat's Principle and Lagrangian Formalism

The eikonal equation and the light ray equation describe the ***local*** behavior of wavefronts and ray trajectories, respectively. In many cases it is, however, convenient to consider the corresponding integral (global) properties.

Let us first consider ***Lagrange's integral invariant theorem***, according to which the integral of vector $n\vec{\tau}$, as well as that of wave vector $\vec{k} = k_0 n\vec{\tau}$, between any two points P_1, P_2 of the optical field is independent of the path traveled, that is,

(21) $$\int_{P_1}^{P_2} n\vec{\tau}\cdot d\vec{r} = \phi(P_2) - \phi(P_1) \quad ,$$

since, according to the eikonal equation (9), $n\vec{\tau}\cdot d\vec{r} = \nabla\phi\cdot d\vec{r} = d\phi$.

Otherwise, if we integrate local equation (10) over ***any given*** surface Σ_C with a closed boundary **C**, and apply Stokes' theorem, we get the ***global law***

(22) $$\iint_{\Sigma_c} \nabla\times(n\vec{\tau})\cdot d\vec{A} = \oint_C n\vec{\tau}\cdot d\vec{r} = 0 \quad .$$

The integral theorem described above is also valid if the integrating contour



should intersect one or more surfaces of refractive index discontinuity. Obviously, these surfaces must be treated as relatively fast but continuous transition areas of the refractive index, across which the local eikonal equation retains its validity. As an application, let us consider such a separation surface intersected by a given closed boundary $C = C_1 + C_2 + C_{12}$, where boundaries C_1 and $C_2$are situated on either side of the separating surface Σ, and the infinitesimal boundary C_{12} effectively intersects this surface (Fig. 6). According to equation (22), we have

(23) $$\oint_C n\vec{\tau} \cdot d\vec{r} = \oint_{C_1} n_1\vec{\tau}_1 \cdot d\vec{r}_1 + \oint_{C_2} n_2\vec{\tau}_2 \cdot d\vec{r}_2 + \oint_{C_{12}} (n_1\vec{\tau}_1 - n_2\vec{\tau}_2) d\vec{r}_2 = 0 \quad .$$

But the integrals over the closed boundaries C_1 and C_2 are null, so that we can deduce the property

(24′) $$(n_1\vec{\tau}_1 - n_2\vec{\tau}_2) \cdot d\vec{r}_2 = 0 \quad ,$$

or the equivalent

(24″) $$(\vec{k}_1 - \vec{k}_2) \cdot d\vec{r}_2 = 0 \quad ,$$

which applies to every point through which the light rays pass across the discontinuity surface Σ. Since the path component $d\vec{r}_2$ in equations (24′) and (24″) represents any infinitesimal portion traversed across surface Σ, these equations are equivalent to the continuity condition of the tangential components of vectors $n\vec{t}$ and $\vec{k}$, respectively. Because it measures the angle of incidence θ_1 and that of refraction θ_2 to the surface normal at the point of ray incidence (Fig. 7), this condition is synonymous with ***the Snell-Descartes law of refraction***

(25) $$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad .$$

Let us now apply Lagrange's integral invariant theorem, that is, equation (21), for the situation in which the integration contour is the trajectory of a light ray itself, so that $n\vec{\tau} \cdot d\vec{r} = n\vec{\tau} \cdot \vec{\tau} ds = n \cdot ds$. In this case, integral (21) between any two points P_1, P_2 on the ray, with the notation $|P_1P_2|$, is termed ***optical path***, and has the following equivalent expressions:

(26) $$[P_1P_2] \stackrel{def}{=} \int_{P_1}^{P_2} n ds = \lambda_0 \int_{P_1}^{P_2} \frac{ds}{\lambda} = c(t_2 - t_1) = \phi_2 - \phi_1 \quad ,$$

in which we've used the expression $n = k/k_0 = \lambda_0/\lambda$ and the equation $nds = cdt$, that is, equation (12). In other words, ***the optical path between two points of a light ray is proportional to the number of wavelengths (since $ds/\lambda = dN$), the time of light travel, and the phase difference between harmonic oscillations of the optical field at the points considered, respectively.***

The concept of optical path allows us to formulate the following general property, called the ***principle of equal optical paths*** (or ***the Malus-Dupin theorem***), according to which, no matter the optical media and the discontinuity surfaces traversed, ***the optical path between any two given wavefronts is the same for all light rays***. The validity of this assertion results from applying equation (26) to all the rays

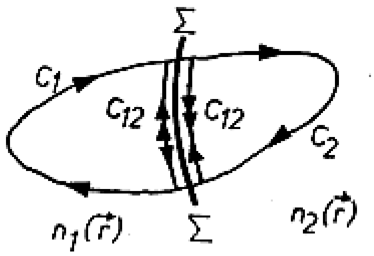
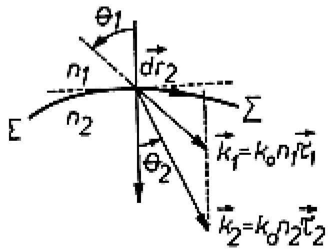


Fig. 6 A separating surface between two different optical media (Σ) and a random closed integration contour ($C = C_1 + C_2 + C_{12}$).



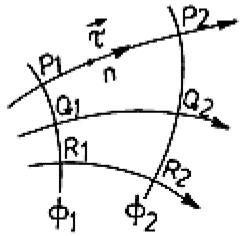


Fig. 7 Light refraction and its fundamental law (geometrical interpretation).

Fig. 8 The principle of equal optical paths (the Malus-Dupin theorem).

of the light beam considered, meaning (see Fig. 8)

(27) $$[P_1P_2] = [Q_1Q_2] = [R_1R_2] = \ldots = \phi_2 - \phi_1 \quad .$$

This principle automatically satisfies the Snell-Descartes law of refraction at the discontinuity surfaces. Let us consider a narrow beam of light rays bound between neighboring rays P_1PP_2 and Q_1QQ_2 passing through surface Σ, separating two homogenous media n_1 and n_2 (see Fig. 9). In virtue of the principle of equal optical paths, we have

$$[P_1PP_2] = [Q_1QQ_2] \quad ,$$

in other words,

$$\underline{n_1 \cdot P_1P'} + n_1 \cdot P'P + \underline{\underline{n_2 \cdot PP_2}} = \underline{n_1 \cdot Q_1Q} + n_2 \cdot QQ' + \underline{\underline{n_2 \cdot Q'Q_2}}$$

where, in construction, the underlined elements compensate each other. Since $P'P = PQ \cdot \sin\theta_1$ and $QQ' = PQ \cdot \sin\theta_2$, this last equation yields the law of refraction $n_1 \sin\theta_1 = n_2 \sin\theta_2$, that is, equation (25).

The principle of equal optical paths also justifies Huygens' construction of successive wavefronts starting from one of them. Thus, as in the previous example, let us consider a separation surface Σ between two homogenous media n_1 and n_2 (see Fig. 10). Based on the given wavefront ϕ_1 in the medium n_1, we must construct geometrically wavefront ϕ_2 in medium n_2, which is separated from wavefront ϕ_1 by the optical path $\phi_2 - \phi_1 = const$. To this end, at the different points P_1, Q_1, R_1, ... of wavefront ϕ_1 we build the normals (light rays) that intersect separating surface Σ in the corresponding points P, Q, R, ... Then we draw the spheres S_P, S_Q, S_R, ... with their centers in said points (P, Q, R, ...) and rays s_{2p}, s_{2q}, s_{2r}, ... given by the condition of equal optical paths

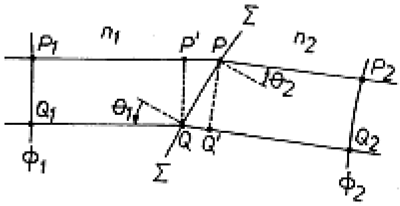


Fig. 9 Deducing law of refraction (25) from the Malus-Dupin theorem.

$$n_1 s_{1p} + n_2 s_{2p} = n_1 s_{1q} + n_2 s_{2q} = n_1 s_{1r} + n_2 s_{2r} = \ldots = \phi_2 - \phi_1 \qquad .$$

Obviously, the envelope of these spheres represent the intended wavefront ϕ_2, and the tangent points P_2, Q_2, R_2, ... are at the same time the points of intersection between light rays P_1PP_2, Q_1QQ_2, R_1RR_2, ... and said wavefront.

Next we will see that the equations of geometrical optics may be deduced from a single variational principle (***Fermat's principle***). Let us thus consider two trajectories (Fig. 11) passing through the same points P_1, and P_2, namely a ***real trajectory***, effectively chosen by the light ray, and a neighboring ***virtual trajectory***, which the light ray does not actually traverse. Obviously, there is an infinity of virtual trajectories neighboring a given real light ray. The variation of the optical path between the two considered trajectories is expressed as

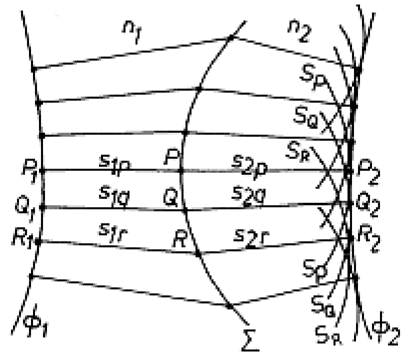


Fig. 10 Justifying Huygens' construction based on the Malus-Dupin theorem.

(28)
$$\delta \int_{P_1}^{P_2} n ds = \int_{P_1}^{P_2} (\delta n) ds + \int_{P_1}^{P_2} n \delta(ds) \qquad .$$

Since the trajectories are close together,

(29)
$$\delta n = \delta \vec{r} \cdot \nabla n \qquad .$$

We also have the successive identities $(ds)^2 = (d\vec{r})^2$, $\delta(ds)^2 = \delta(d\vec{r})^2$, so $ds \cdot \delta(ds) = d\vec{r} \cdot \delta(d\vec{r})$, or, considering that operators d and δ are commutable,

(30) $$\delta(ds) = \frac{d\vec{r}}{ds} \cdot \delta(d\vec{r}) = \vec{\tau} \cdot d(\delta\vec{r}) \quad .$$

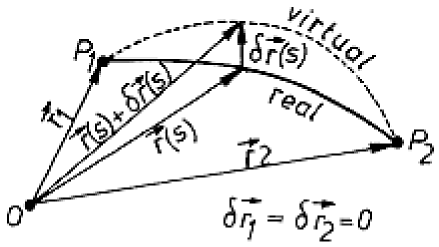


Fig. 11 A real luminous trajectory and a neighboring virtual trajectory (formulating Fermat's principle).

By introducing expressions (29) and (30) into equation (28), we get

(31) $$\delta \int_{P_1}^{P_2} n ds = \int_{P_1}^{P_2} (\delta\vec{r} \cdot \nabla n) ds + \int_{P_1}^{P_2} n\vec{\tau} \cdot d(\delta\vec{r}) \quad .$$

If we integrate by parts the second integral of the right component, we get

(32) $$\int_{P_1}^{P_2} n\vec{\tau} \cdot d(\delta\vec{r}) = n\vec{\tau} \cdot \delta\vec{r} \Big|_{P_1}^{P_2} - \int_{P_1}^{P_2} \delta\vec{r} \cdot d(n\vec{\tau}) \quad ,$$

so that, in the end, equation (31) is written as

(33) $$\delta \int_{P_1}^{P_2} n ds = n_2\vec{\tau}_2 \cdot \delta\vec{r}_2 - n_1\vec{\tau}_1 \cdot \delta\vec{r}_1 + \int_{P_1}^{P_2} \left[\nabla n - \frac{d}{ds}(n\vec{\tau}) \right] \cdot \delta\vec{r}(s) ds \quad .$$

But points P_1, P_2 at the two ends are supposed to be fixed in position, that is, $\delta\vec{r}_1 = \delta\vec{r}_2 = 0$, and the variation $\delta\vec{r}(s)$ is arbitrary. It thus follows that the light ray equation $\frac{d}{ds}(n\vec{\tau}) = \nabla n$ and, implicitly, the eikonal equation, are mathematically equivalent to the variational formulation

(34) $$\delta \int_{P_1}^{P_2} n ds = 0 \quad .$$

This yields ***Fermat's principle*** (1657), according to which ***the real trajectory of a light ray that links any two given points P_1, P_2 is determined by the condition that the corresponding optical path should be stationary (extremal from the perspective of variational calculus)***, meaning

(35) $$\int_{P_1}^{P_2} n ds = \text{ stationary (extremal),}$$

where stationary (extremal) means minima, maxima, or constant. In other words, the real trajectory of the light ray represents an ***extremal trajectory of the optical path***. Obviously, this trajectory remains the same no matter the direction of light propagation (the property of the **reversibility of light rays**). In homogenous media (n = constant), light propagates especially along the extremal geometrical path

(36) $$\int_{P_1}^{P_2} ds = \text{constant} = \text{stationary (extremal)} \quad ,$$

that is, in a straight line (minimum).

Note that, from a historical perspective, geometrical optics has developed as a theory of light rays, directly defined by Fermat's principle, that is, by the principle of trajectories along which the optical path is stationary (extremal). The first success of Fermat's principle was, of course, the deduction of the already known laws of reflection and refraction. Let us also deduce in this way the law of refraction across a surface Σ, separating two homogenous media n_1, n_2 (see Fig. 12). According to said principle, along the real trajectory connecting two given points P_1, P_2, we have

(37) $$n_1 s_1 + n_2 s_2 = \text{ stationary (extremal).}$$

Upon a virtual displacement $d\vec{s}_1$ of the point of incidence of the light ray on surface Σ, it thus follows that

(38) $$n_1 \delta s_1 + n_2 \delta s_2 = 0 \quad .$$

But $s^2 = \vec{s}_2$, $s\delta s = \vec{s} \cdot d\vec{s}$, $\delta s = \frac{\vec{s}}{s} d\vec{s} = \vec{\tau} \cdot d\vec{s}$, and $\overrightarrow{P_1 P_2} = \vec{s}_1 + \vec{s}_2 = \overrightarrow{constant}$, meaning

$$d\vec{s}_1 = -d\vec{s}_2$$

so that

(39) $$(n_1 \vec{\tau}_1 - n_2 \vec{\tau}_2) \cdot d\vec{s}_1 = 0$$

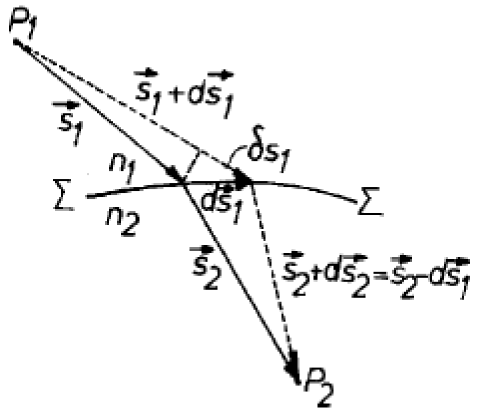


Fig. 12 Deducing the laws of refraction in Fermat's principle

Since the virtual displacement across the surface is arbitrary, equation (39) is equivalent to the condition of continuity of the tangential component of the vector, that is, to the ***Snell-Descartes law of refraction***. The law of reflection is similarly deduced if we consider the points located within the same medium.

The method of deducing natural laws from an integral variational principle, expressed for the first time through Fermat's principle in geometrical optics, has



proven much more general, and has dominated the entirety of subsequent development of physical theories. For example, take Newton's second law

(40) $$m\frac{d\vec{v}}{dt} = -\nabla U \quad ,$$

through which classical mechanics describes the motion of a material point of mass m and speed $\vec{v}$ in a field of force $\vec{F} = -\nabla U$, determined by the potential energy $U(\vec{r})$. From the law of the conservation of energy

(41) $$\frac{1}{2}mv^2 + U(\vec{r}) = E \quad ,$$

in which E is the total energy, through the gradient operation, we get

(42) $$mv \cdot \nabla v = -\nabla U \quad ,$$

so that equation (40) can also be written

(43) $$\frac{1}{v}\frac{d\vec{v}}{dt} = \nabla v \quad .$$

Taking into account the fact that $ds = vdt$, and introducing the unit vector $\vec{\tau} = \frac{\vec{v}}{v}$ of the tangent to the trajectory, from equation (43) we get the ***equation of particle trajectory***, written as

(44) $$\frac{d}{ds}(v\vec{\tau}) = \nabla v \quad .$$

This equation is the analogue in classical mechanics of the light ray equation, equation (13), the refractive index $n(\vec{r}) = c/v(\vec{r})$ being here replaced by the particle's speed $v(\vec{r}) = \left[(2/m)\cdot\left(E - U(\vec{r})\right)\right]^{1/2}$. Accordingly, the analogue of Fermat's principle, equation (35), is written as

(45) $$\int_{P_1}^{P_2} vds \quad = \quad \text{stationary (extremal)},$$

and is known as the ***Maupertuis-Euler principle*** (1744). This is how the ***opto-mechanical analogy*** arose, between the matter of tracing light rays in a medium of refractive index $n(\vec{r})$ and that of determining the trajectories of particles in field of forces described by the function of potential energy $U(\vec{r})$. The foundations of this

analogy were further strengthened by Hamilton, who applied variational calculus to the integral of the optical path in equation (35) in the field of geometrical optics (***Theory of systems of rays***, 1828 – 1837), as well as to the integral of "action" in equation (45) in the field of classical dynamics (***On the application to dynamics of a general mathematical method previously applied to optics***, 1834).

For its beauty and comprehensiveness, we will now present the Lagrangian and Hamiltonian formulation of geometrical optics. For convenience, we will reiterate the light ray trajectory in a Cartesian system of coordinates, going from the parametrical representation $x(s)$, $y(s)$, $z(s)$ to the representation $x(z)$, $y(z)$, z, depending on the independent variable z (Fig. 13). Thus, the length component of the trajectory is written as

(46) $$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2} = (1 + x'^2 + y'^2)^{1/2} dz \quad ,$$

where

(47) $$\begin{cases} x' = \dfrac{dx}{dz} = \dfrac{\tau_x}{\tau_z} \,, \\ y' = \dfrac{dy}{dz} = \dfrac{\tau_y}{\tau_z} \,, \end{cases}$$

and

(48) $$\begin{cases} \tau_x = \dfrac{dx}{ds} = \cos\alpha \,, \\ \tau_y = \dfrac{dy}{ds} = \cos\beta \,, \\ \tau_z = \dfrac{dz}{ds} = \cos\gamma \,, \end{cases}$$

are the components of unit vector $\vec{\tau} = \frac{d\vec{r}}{ds}$ (the direction cosines of the tangent to the trajectory). We will change the integration variable for the optical path from s to z, that is,

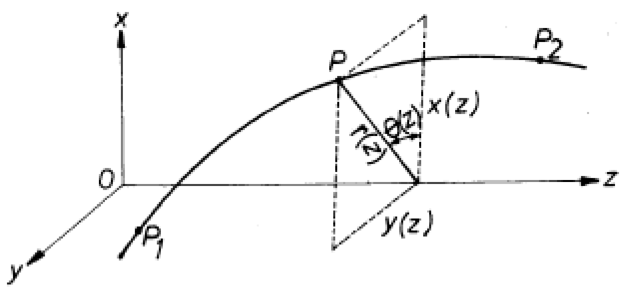


Fig. 13 Representation $(x(z), y(z), z)$ of the light ray trajectory.



(49)
$$[P_1P_2] = \int_{P_1}^{P_2} n(s)ds = \int_{z_1}^{z_2} n\left(\frac{ds}{dz}\right)dz =$$
$$= \int_{z_1}^{z_2} L[x(z), y(z), x'(z), y'(z), z] \cdot dz \ ,$$

where

(50)
$$L(x, y, x', y', z) = n\frac{ds}{dz} = \frac{n}{\tau_z} = \frac{n}{\cos\gamma} = n(x, y, z) \cdot (1 + x'^{\,2} + y'^{\,2})^{1/2} \ ,$$

is the ***optical Lagrangian***. According to Fermat's principle, the real light ray trajectory must satisfy equation (34), that is,

(51)
$$\delta \int_{z_1}^{z_2} L(x, y, x', y', z)dz = 0 \ .$$

As is shown in variational calculus, the necessary conditions imposed by equation (51) are given by the ***Euler-Lagrange equations***

(52)
$$\begin{cases} \dfrac{d}{dz}\left(\dfrac{\partial L}{\partial x'}\right) = \dfrac{\partial L}{\partial x}, \\ \dfrac{d}{dz}\left(\dfrac{\partial L}{\partial y'}\right) = \dfrac{\partial L}{\partial y}. \end{cases}$$

These equations are actually the light ray equation, namely equation (13)

$$\frac{d}{ds}(n\vec{\tau}) = \nabla n \ ,$$

or, split into components,

(53)
$$\begin{cases} \dfrac{d}{ds}(n\tau_x) = \dfrac{\partial n}{\partial x}, \\ \dfrac{d}{ds}(n\tau_y) = \dfrac{\partial n}{\partial y}, \\ \dfrac{d}{ds}(n\tau_z) = \dfrac{\partial n}{\partial z}. \end{cases}$$

Indeed, by deriving the Lagrangian expression, namely equation (50), we get

(54)
$$\frac{\partial L}{\partial x'} = \frac{nx'}{(1 + x'^{\,2} + y'^{\,2})^{1/2}} = n\frac{dx}{ds} = n\tau_x \ .$$

The first equation (52) is explicitly written as

$$\frac{d}{dz}(n\tau_x) = \left(1 + x'^2 + y'^2\right)^{1/2} \cdot \frac{\partial n}{\partial x} \qquad \text{or} \qquad \frac{d}{ds}(n\tau_x) = \frac{\partial n}{\partial x} \qquad ,$$

meaning it is equation (53). Similarly, the second (52) equation is the second (53) equation.

Note that only the first two (53) equations are independent, while the third equation follows automatically from the other two and from the purely geometrical equation

(55) $$\vec{\tau}^2 = \tau_x^2 + \tau_y^2 + \tau_z^2 = 1 \quad .$$

Indeed, by multiplying equation (55) by dn/ds, and by deriving it over s and multiplying it by n, respectively, we get

$$\left(\tau_x^2 + \tau_y^2 + \tau_z^2\right) \cdot \frac{dn}{ds} = \frac{\partial n}{\partial x} \cdot \tau_x + \frac{\partial n}{\partial y} \cdot \tau_y + \frac{\partial n}{\partial z} \cdot \tau_z \quad ,$$

$$\left(\tau_x \frac{d\tau_x}{ds} + \tau_y \frac{d\tau_y}{ds} + \tau_z \frac{d\tau_z}{ds}\right) n = 0 \quad ,$$

from which, through addition, we get

(56) $$\tau_x\left[\frac{d}{ds}(n\tau_x) - \frac{\partial n}{\partial x}\right] + \tau_y\left[\frac{d}{ds}(n\tau_y) - \frac{\partial n}{\partial y}\right] + \tau_z\left[\frac{d}{ds}(n\tau_z) - \frac{\partial n}{\partial z}\right] = 0 \quad .$$

Obviously, the third (53) equation is an identity that brings nothing new beyond the first two.

From the Lagrangian formalism presented above, through equations (50) – (52), we can pass into Hamiltonian formalism by defining the ***canonical optical momenta***:

(57) $$\begin{cases} p_x = \dfrac{\partial L}{\partial x'} = n\tau_x \, , \\ p_y = \dfrac{\partial L}{\partial y'} = n\tau_y \, , \end{cases}$$

and the ***optical Hamiltonian***

(58) $$H = p_x x' + p_y y' - L(x, y, x', y', z) = p_x \frac{\tau_x}{\tau_z} + p_y \frac{\tau_y}{\tau_z} - \frac{n}{\tau_z} \quad .$$

By replacing the direction cosines with momenta, based on equations (57) and (55), we obtain the expression of the Hamiltonian dependent on the canonical conjugate variables (x, p_x), (y, p_y) and the independent variable z, in the form

(59) $$H(x, y, p_x, p_y, z) = -\left[n^2(x, y, z) - \left(p_x^2 + p_y^2\right)\right]^{1/2} = -n\tau_z = -n\cos\gamma \quad .$$

Note that, while the optical coordinates (x, y) can have any value, the range of optical momenta (p_x, p_y) is limited by the condition

$$p_x^2 + p_y^2 = n^2(\tau_x^2 + \tau_y^2) = n^2(1 - \tau_z^2) = (n\sin\gamma)^2 \leq n^2 \quad .$$

Correspondingly, we have $|H| \leq n$.

The total differential of the optical Hamiltonian (59) as a function of coordinates and momenta is

(60) $$dH = \frac{\partial H}{\partial x}dx + \frac{\partial H}{\partial y}dy + \frac{\partial H}{\partial p_x}dp_x + \frac{\partial H}{\partial p_y}dp_y + \frac{\partial H}{\partial z}dz \quad .$$

On the other hand, from the defining equation (58), it follows that

(61) $$dH = x'dp_x + \underline{p_x dx'} + y'dp_y + \underline{\underline{p_y dy'}} - \frac{\partial L}{\partial x}dx \quad - \frac{\partial L}{\partial y}dy - \underline{\frac{\partial L}{\partial x'}dx'} - \underline{\underline{\frac{\partial L}{\partial y'}dy'}} - \frac{\partial L}{\partial z}dz \quad ,$$

where the underlined terms compensate each other based on the definition of momenta itself, equation (57). Moreover, according to the Euler-Lagrange equations (52), we have

(62) $$\frac{\partial L}{\partial x} = \frac{dp_x}{dz}, \qquad \frac{\partial L}{\partial y} = \frac{dp_y}{dz} \quad ,$$

so that equation (61) is rewritten as

(63) $$dH = \frac{dx}{dz}dp_x + \frac{dy}{dz}dp_y - \frac{dp_x}{dz}dx - \frac{dp_y}{dz}dy - \frac{\partial L}{\partial z}dz \quad .$$

By identifying the two expressions (60), (63) of the total differential dH, we finally get the differential equations for the canonical variables, called ***canonical equations***, or ***Hamilton's equations***

(64) $$\begin{cases} \dfrac{dx}{dz} = \dfrac{\partial H}{\partial p_x}, & \dfrac{dp_x}{dz} = -\dfrac{\partial H}{\partial x}, \\ \dfrac{dy}{dz} = \dfrac{\partial H}{\partial p_y}, & \dfrac{dp_y}{dz} = -\dfrac{\partial H}{\partial y}, \end{cases}$$

as well as $\partial H/\partial z = -\partial L/\partial z$. Instead of two second order Euler-Lagrange equations,

we've thus obtained four first order Hamilton equations. By knowing the system's Hamiltonian, equation (59), and by specifying the boundary conditions in a point $P_0(z_0)$, integrating the (64) equations allows us to determine the state of the light ray in any other point $P(z)$, that is, its x, y position and direction of propagation $p_x = n\tau_x$, $p_y = n\tau_y$. When interpreting geometrically, it is convenient to consider the canonical variables in ***phase space*** (x, y, p_x, p_y).

Using form (59) of the Hamiltonian, the reader can easily verify that the canonical equations (64) lead to definitions (57) and to the first two equations of set (53). This means that the entirety of Hamiltonian equations is absolutely equivalent to the Euler-Lagrange equations.

The study of luminous trajectories can equally be elaborated using the ***Hamilton-Jacobi method***. By defining the ***optical action*** $S(x(z), y(z), z)$ using the equation

(65) $$S = \int_{P_1}^{P} L dz = \int_{P_1}^{P} (p_x x' + p_y y' - H) dz \quad ,$$

known to us from analytical mechanics, in which P_1 is a fixed point, and P is an arbitrary point along the real luminous trajectory, just as in analytical mechanics, we obtain

(66) $$\frac{\partial S}{\partial z} = -H, \ \frac{\partial S}{\partial x} = p_x, \ \frac{\partial S}{\partial y} = p_y \quad .$$

Since the Hamiltonian H is a function of the canonical conjugate variables (x, p_x), (y, p_y), and of z, the first (66) equation becomes

(67) $$\frac{\partial S}{\partial z} + H(x, y; \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}; z) = 0 \quad .$$

We've thus obtained the **Hamilton-Jacobi equation** for optical action. Using the effective form (59) of the optical Hamiltonian, based on equation (67) we get

(68) $$\frac{\partial S}{\partial z} - \left[n^2(x, y, z) - \left(\frac{\partial S}{\partial x}\right)^2 - \left(\frac{\partial S}{\partial y}\right)^2 \right]^{1/2} = 0 \quad ,$$

which can readily be written as

(68′) $$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2(x, y, z) \quad .$$

By comparing equation (68′) to equation (6), we reach the conclusion that the eikonal

is identical to the optical action, and the eikonal equation is in fact the Hamilton-Jacobi equation. Thus, the "circle" of opto-mechanical analogy is complete.

This allows us to affirm that between geometrical optics and analytical mechanics there is a perfect analogy. In quantum mechanics, the Hamilton-Jacobi equation (that is, the eikonal equation) is a limit case (for $h \to 0$, h = Planck's constant) of Schrödinger's equation, fundamental to the entirety of nonrelativistic quantum mechanics. Therefore, analytical mechanics and optics may be considered, in the sense of the ***correspondence principle***, to be particular cases of quantum mechanics. Light's dual nature (wavelike and corpuscular-photonic) is thus integrated into the dual nature of quantum micro-objects.

1.3 General Conditions of Stigmatism

Let us consider a conical (homocentric) light beam emitted by a punctual source P_1 (Fig. 14.a). Usually, out of the infinity of rays within the beam, only one will pass through another point P_2, namely the extremal trajectory that satisfies Fermat's principle. On the other hand, the ideal function of image forming optical systems is directing the light beam in such a way that every point P_1 in the object space has a single corresponding point P_2 within the image space. For this reason, we will now deal with those exceptional cases in which points P_1 and P_2 are linked by an infinity of rays (Fig. 14.b).

Stigmatism is the fundamental concept of the geometrical theory of optical images. The term comes from Greek στίγμα, which means point. An optical system is stigmatic, or punctual, for the pair of points P_1, P_2 if a conical light beam with its apex at P_1 is transformed into a conical light beam with its apex at in P_2. Point P_2 is called the ***stigmatic image*** of point P_1. Obviously, if we inverse the direction of the light rays, point P_1 becomes the stigmatic image of point P_2. The object and image points thus defined form a ***pair of stigmatic, or conjugate, points*** of the optical system considered. Depending on whether the light rays actually intersect or intersect through their extensions (rectilinear, in homogenous media), the object or image point is termed a ***real*** or a ***virtual*** point.

Contemporary
Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017

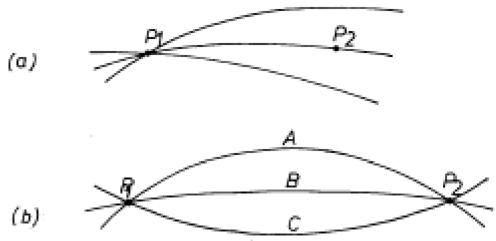


Fig. 14.a,b Defining stigmatism.

Generally, irrespective of the complexity of the shape of the wavefronts, they are always spherical in the immediate vicinity of the conjugate points, into which they degenerate. By extending the principle of equal optical paths, we see that the optical path, the time of light propagation, the number of wavelengths and the phase difference between the two conjugate points P_1, P_2 are the same for all light rays A, B, C, ... that pass through these points (Fig. 14.b). The **stigmatism condition** for conjugate points P_1, P_2 is written as

(69) $$[P_1P_2] = [P_1AP_2] = [P_1BP_2] = [P_1CP_2] = ... = \quad \text{constant}$$

and shows the only way in which light may propagate between two points by effectively and simultaneously taking multiple neighboring paths. Indeed, only in this way is the condition of stationary optical path between conjugate points, imposed by Fermat's principle, equation (35), satisfied, for any light ray within the beam considered. The property of equal time of light propagation between conjugate points is termed ***tautochronism***.

Perhaps the physical significance of the notion of image in geometrical optics most clearly emerges from the property of equal wavelength number, and from the property according to which the relative phase of the harmonic waves that propagate along the various rays is the same at the conjugate points. In order to illustrate how a perfect image is formed, in Fig. 15.a,b,c we have the reconstruction of spherical waves at a ***Cartesian surface of refraction*** (***Cartesian oval***), defined as the surface Σ separating two homogenous media n_1, n_2, whose points I satisfy the stigmatism condition (only) for a given pair of conjugate points P_1, P_2. Generally, the Cartesian oval is an ***aspheric*** bipolar surface, with rotational symmetry around the axis passing through the considered conjugate points. Therefore, when both conjugate points are real, that is, they are points through which the light rays effectively pass (Fig. 15.a), the Cartesian surface satisfies the equation

(70) $$[P_1IP_2] = n_1 \cdot P_1I + n_2 \cdot IP_2 = \lambda_0\left(\frac{P_1I}{\lambda_1} + \frac{IP_2}{\lambda_2}\right) \quad = \text{constant.}$$

In other words, irrespective of the point of incidence I on the surface, the

conjugate points are separated by the same number of wavelengths (in Fig. 15, the number is 23).

The stigmatism condition (70) can be extended to the cases in which one or both conjugate points are virtual. For example, let us consider the case in which P_1 is real and P_2 is virtual (Fig. 15.b). The rectilinear extensions of light rays in medium n_2 now pass through P_2. According to the principle of equal optical paths, between the real point P_1 and a given wavefront ϕ_2 in medium n_2 we have

$$[P_1IJ] = n_1 \cdot P_1I + n_2 \cdot IJ = n_1 \cdot P_1I + n_2 \cdot (P_2J - P_2I) \quad = \quad \text{constant}$$

irrespective of the point of incidence I. Segment P_2J is the ray of considered spherical surface ϕ_2, and is constant, so that we obtain the stigmatism condition (the equation for the Cartesian surface), written as

(71) $$n_1 \cdot P_1I - n_2 \cdot P_2I = \lambda_0\left(\frac{P_1I}{\lambda_1} - \frac{P_2I}{\lambda_2}\right) = \quad \text{constant.}$$

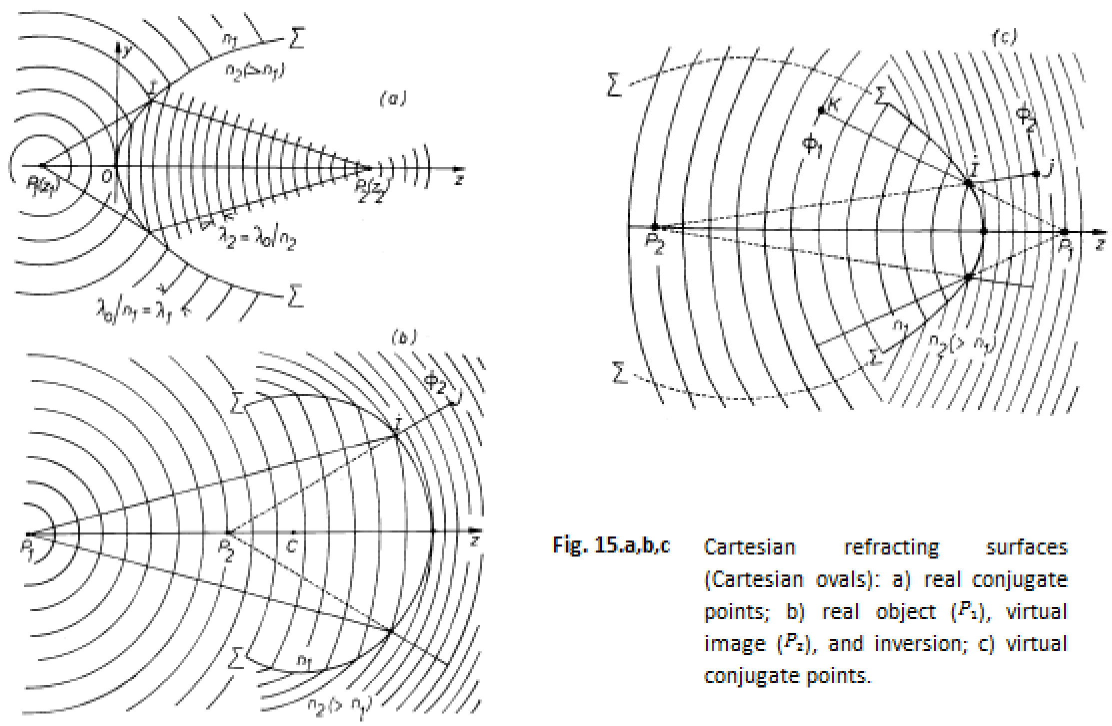


Fig. 15.a,b,c Cartesian refracting surfaces (Cartesian ovals): a) real conjugate points; b) real object (P_1), virtual image (P_2), and inversion; c) virtual conjugate points.

This time, unlike condition (70), irrespective of the position of the point of incidence I,

the conjugate points are separated by the same wavelength number difference between the real ray P_1I and virtual ray P_2I (in Fig. 15.b, 0 was chosen as the value of the constant in equation (71)). In a similar fashion, it can be demonstrated that the Cartesian surface equation in the situation in which point P_1 is virtual and P_2 is real takes the form

(72) $$-n_1 \cdot IP_1 + n_2 \cdot IP_2 = \lambda_0\left(-\frac{IP_1}{\lambda_1} + \frac{IP_2}{\lambda_2}\right) = \quad \text{constant.}$$

Note that in the cases described by equations (71) and (72), the conjugate points are located on the same side of the Cartesian surface. In the special case when the difference of optical paths between the real and virtual ray is null, the Cartesian surface degenerates into a spherical surface (see Fig. 15.b) and the corresponding conjugate points are termed ***Weierstrass points*** (or ***Young's points***). These points have significant practical importance, since, on the one hand, spherical surfaces are easiest to polish, and, on the other hand, Weierstrass points are not only stigmatic, but also aplanatic (see section 2.1). Finally, if both conjugate points are virtual (Fig. 15.c), by applying the principle of equal paths between a wavefront ϕ_1 in medium n_1 and a wavefront ϕ_2 in medium n_2, we get

$$[KIJ] = n_1 \cdot KI + n_2 \cdot IJ = n_1 \cdot (KP_1 - IP_1) + n_2 \cdot (P_2J - P_2I) = \quad \text{constant,}$$

irrespective of the position of the point of incidence I. But curvature radii KP_1 and P_2J of spherical wavefronts ϕ_1 and ϕ_2, respectively, are constant, so we can write

(73) $$-n_1 \cdot IP_1 - n_2 \cdot P_2I = -\lambda_0\left(\frac{IP_1}{\lambda_1} + \frac{P_2I}{\lambda_2}\right) = \text{constant.}$$

Therefore, just as in the case of real conjugate points, equation (70), virtual conjugate points are situated on either side of the Cartesian surface, and, irrespective of the position of the point of incidence I, they are separated by the same number of wavelengths (in Fig. 15.c, the number is 18).

Let us go over the results obtained in equations (70) – (73). We have:

(74)
$$\begin{array}{lrl}
P_1, P_2 \text{ real} & n_1 \cdot P_1I + n_2 \cdot IP_2 = & \text{constant,} \\
P_1, P_2 \text{ virtual} & -n_1 \cdot IP_1 - n_2 \cdot P_2I = & \text{constant,} \\
P_1 \text{ real}, P_2 \text{ virtual} & n_1 \cdot P_1I - n_2 \cdot P_2I = & \text{constant,} \\
P_1 \text{ virtual}, P_2 \text{ real} & -n_1 \cdot IP_1 + n_2 \cdot IP_2 = & \text{constant,}
\end{array}$$

where all segments are considered positive. In short, the ***conditions of rigorous***

stigmatism, that is, equations (74), are written as

(75) $$[P_1IP_2] = n_1 \cdot P_1I + n_2 \cdot IP_2 = \quad \text{constant},$$

where the segments are considered from an algebraic perspective, meaning we ***accept the convention that the optical path is positive if traversed along the direction of propagation, and negative if traversed in the opposite direction.***

In fact, we distinguish between two categories of situations, namely

(76) $$n_1 \cdot |P_1I| \pm n_2 \cdot |IP_2| = \quad \text{constant},$$

where the plus sign pertains to the case in which the conjugate points are situated on different sides of the Cartesian surface, and the minus sign, to that in which they are located on the same side.

Note that, from a formal perspective, we can specialize the formulas presented above for reflection by simply replacing n_2 with $-n_1$, so that equation (76) becomes

(77) $$|P_1I| \mp |IP_2| = \quad \text{constant},$$

where this time the minus sign pertains to the case in which the conjugate points are located on either side of the mirror, and the plus sign, to the case in which they are located on the same side. Indeed, upon studying Fig. 16 we reach the conclusion that ***the Cartesian surfaces of reflection*** are either hyperboloids of revolution (Fig. 16.b,c, with focal points P_1 and P_2, one real, one virtual), or ellipsoids of revolution (Fig. 16.e,f, with focal points P_1 and P_2, both real or both virtual). A special case of hyperbolic mirror is the plane mirror (Fig. 16.d, when the constant in equation (77) is null). Also, when one of the focal points moves to infinity, the elliptical mirror becomes parabolic (Fig. 16.g,h).

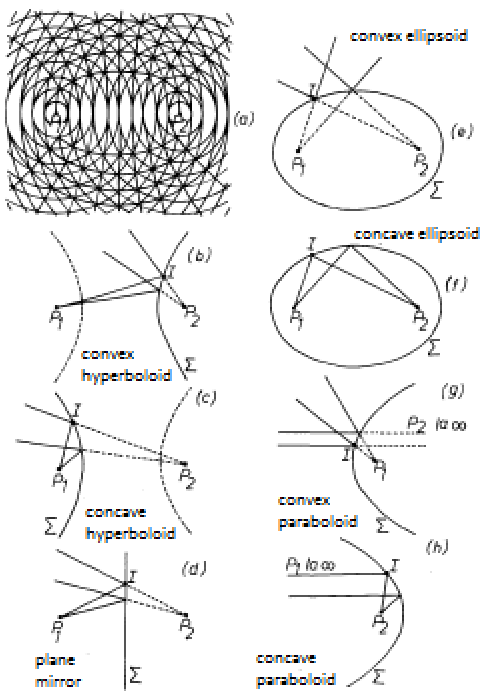


Fig. 16 Cartesian reflecting surfaces; eillipsoids of revolution (a, e, f), hyberboloids of revolution (b, c, and, as a special case, d), and paraboloids of revolution (g, h).

Generally, unlike the Cartesian surfaces of reflection of equation (77), which have conical sections, the Cartesian surfaces of refraction, equation (76), are much more complex. Thus, if we fix the conjugate points $P_1(z_1)$ and $P_2(z_2)$, and choose a system of Cartesian coordinates yOz, with the origin in vertex O of surface Σ (Fig. 15.a), equation (76) becomes

(78) $$n_1 \cdot \sqrt{(z-z_1)^2+y^2} \pm n_2 \cdot \sqrt{(z-z_2)^2+y^2} = n_1 \cdot |z_1| \pm n_2 \cdot |z_2| \quad ,$$

where the plus sign pertains to the case in which points P_1 and P_2 are located on different sides of the origin O, that is, $z_1 z_2 < 0$, while the minus sign, to that in which they are situated on the same side, meaning $z_1 z_2 > 0$. If we eliminate the roots by squaring the equation twice and order the terms according to the decreasing exponents of values y and z, we finally come to the ***Cartesian oval equation***, written as

(79) $$\begin{aligned}&(n_1^2-n_2^2)^2 \cdot (y^2+z^2)^2 - 4(n_1^2-n_2^2)\cdot(n_1^2 z_1 - n_2^2 z_2)\cdot z(y^2+z^2)+\\ &+4n_1 n_2(n_1 z_1 - n_2 z_2)(n_1 z_2 - n_2 z_1)(y^2+z^2) + 4(n_1^2 z_1 - n_2^2 z_2)^2 z^2 -\\ &-8n_1 n_2(n_1-n_2)(n_1 z_1 - n_2 z_2) z_1 z_2 z = 0 \quad .\end{aligned}$$

This equation represents the meridional section of a fourth order surface of revolution. For certain values of parameters n_1, n_2, z_1, z_2, the Cartesian oval degenerates into a second order surface. Thus, the third and fourth order terms eliminate each other if $n_1^2 - n_2^2 = 0$, that is, if $n_2 = \pm n_1$. The situation $n_2 = n_1$ is trivial (adjacent media are identical), and the case in which $n_2 = -n_1$ occurs in reflection, and has been discussed above. Moreover, if the following equation applies:

(80) $$n_1 z_1 = n_2 z_2 \ , \qquad (z_1 z_2 > 0) \ ,$$

then equation (79) becomes

$$(n_1^2-n_2^2)\cdot(y^2+z^2) - 2z(n_1^2 z_1 - n_2^2 z_2) = 0,$$

meaning sphere

(80) $$y^2 + \left(z - \frac{n_1^2 z_1 - n_2^2 z_2}{n_1^2 - n_2^2}\right)^2 = = \left(\frac{n_1^2 z_1 - n_2^2 z_2}{n_1^2 - n_2^2}\right)^2 ,$$

with its center C at point

(82) $$y_c = 0\,, \qquad z_c = \frac{n_1^2 z_1 - n_2^2 z_2}{n_1^2 - n_2^2} \qquad ,$$

and a radius $r = |z_c|$. Using equation (82) as an expression for z_c, the position of the ***Weierstrass points***, defined by equation (80), can also be written as

(83) $$z_1 = \left(1 + \frac{n_2}{n_1}\right) z_c\,,$$
$$z_2 = \left(1 + \frac{n_1}{n_2}\right) z_c\,.$$

Note that Weierstrass points $P_1(z_1)$, $P_2(z_2)$, and the center of curvature $C(z_c)$ are located on the same side of surface Σ, as can be seen in Fig. 15.b (for applications, see Chapter 2.1).

Finally, if one of the conjugate points is located at infinity ($|z_1| \to \infty$ ***or*** $|z_2| \to \infty$), then the Cartesian oval is an ellipsoid or hyperboloid of revolution. In order to demonstrate this, we note that the left member of equation (79) represents a second order polynomial in z_1 (or z_2), so that the oval equation can also be written as

$$\left[(n_1^2 - n_2^2)z^2 - n_2^2 y^2 - 2n_2(n_1 - n_2)z_2 z\right] \cdot z_1^2 + \ldots = 0 \qquad .$$

If $|z_1| \to \infty$, then the brackets in this last equation is eliminated, that is,

(84) $$(n_1^2 - n_2^2)z^2 - n_2^2 y^2 - 2n_2(n_1 - n_2)z_2 z = 0$$

or, written in canonical form,

(85) $$\frac{\left(z - \dfrac{n_2 z_2}{n_1 + n_2}\right)^2}{\left(\dfrac{n_2 z_2}{n_1 + n_2}\right)^2} + \frac{y^2}{\dfrac{n_2 - n_1}{n_1 + n_2} z_2^2} = 1 \qquad .$$

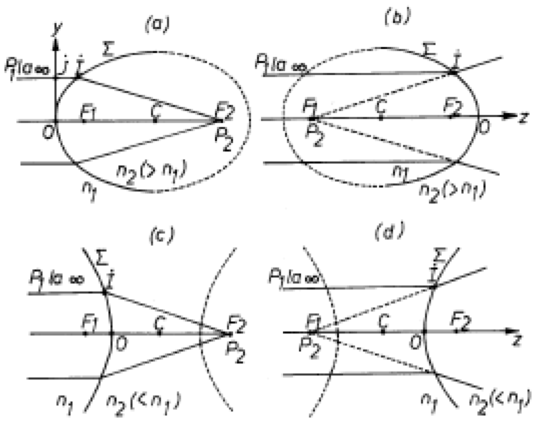


Fig. 17 The Cartesian refracting surfaces when P_1 is located at infinity: ellipsoids of revolution (a, b), hyperboloids of revolution (c, d).

Equation (85), if $n_2 > n_1$, represents an ellipsoid of revolution (Fig. 17.a,b), and if

$n_2 < n_1$, a hyperboloid of revolution (Fig. 17.c,d). We can reach the same result if we note that, when $z_1 \to -\infty$, the stigmatism condition requires that the optical path between an incident wave plane (any such plane) and the image point P_2 should be constant. For example, in the case of Fig. 17.a, we therefore have $[JIP_2] = [OP_2]$, that is,

$$n_1 z + n_2 \cdot \sqrt{(z-z_2)^2 + y^2} = n_2 z_2 \quad ,$$

based on which, by isolating the root, squaring and ordering, we get equation (85). The same result also applies to the other situations presented in Fig. 17.

According to equation (85), the large semi-axis **a**, the small semi-axis **b**, the focal distance **f**, and the eccentricity $e = f/a$ are expressed as

(86) $$a = \frac{n_2|z_2|}{n_1+n_2}, \quad b = \left(\frac{n_2-n_1}{n_2+n_1}\right)^{1/2}|z_2|, \quad f = \sqrt{a^2-b^2} = \frac{n_1|z_2|}{n_1+n_2}, \quad e = \frac{n_1}{n_2} < 1 \quad ,$$

for the ellipsoid ($n_2 > n_1$), and

(87) $$a = \frac{n_2|z_2|}{n_1+n_2}, \quad b = \left(\frac{n_1-n_2}{n_2+n_1}\right)^{1/2}|z_2|, \quad f = \sqrt{a^2+b^2} = \frac{n_1|z_2|}{n_1+n_2}, \quad e = \frac{n_1}{n_2} > 1 \quad ,$$

for the hyperboloid ($n_2 < n_1$). In both cases, the center C of the conical section is located at point

(88) $$z_c = \frac{n_2 z_2}{n_1+n_2}, \quad y_c = 0.$$

Image P_2 coincides with the right focal point (F_2) when $z_2 = z_c + f (> 0)$, or with the left focal point (F_1) when $z_2 = z_c - f (< 0)$.

In the case of reflection ($n_2 = -n_1$), equation (84) becomes

(89) $$y^2 = 4 z_2 z \quad ,$$

and represents a paraboloid of rotation, of parameter $p = 2z_2$. If $p > 0$, the image is virtual (Fig. 16.g), and if $p < 0$, it is real (Fig. 16.h).

The Cartesian surfaces of reflection are important for building telescopes and projectors. Thus, the objective in reflection telescopes (Newton, Herschel, Gregory, Cassegrain) is a concave parabolic mirror, and the secondary mirror is elliptical and concave (Gregory), or hyperbolic and convex (Cassegrain), see section 2.5 (Fig. 61).

Moreover, the property of rigorous stigmatism of Cartesian surfaces of



refraction is used in fashioning ***aspheric lenses***. Generally, if we refer to Fig. 15.a,b,c, an aspheric lens fashioned out of optical medium n_2 is limited by the Cartesian surface Σ and by any spherical surface ϕ_2 with its center at P_2. In practice, Cartesian surfaces with conical section are used, as is illustrated in Fig. 18 in the case of sphero-elliptic (a), plano-hyperbolic (b), or double-hyperbolic (c) lenses. Because of the absence of spherical aberrations (see Chapter 2.8), aspheric lenses can have much larger aperture diameters D and much smaller focal distances f than spherical lenses. Consequently, we can achieve very small (up to 0.6 in practice) ***numbers*** $f \overset{\text{def}}{=} f/D$ (see Chapter 2.6) and a very large luminous flux density within the image plane. Therefore, aspheric lenses allow the most efficient usage of light sources and detectors, which accounts for their numerous applications in present-day optical systems for communications and control. Finally, let us note the use of the properties of Weierstrass points in fashioning stigmatic and aplanatic lenses, and microscope objectives of very high numerical aperture (see Chapter 2.1).

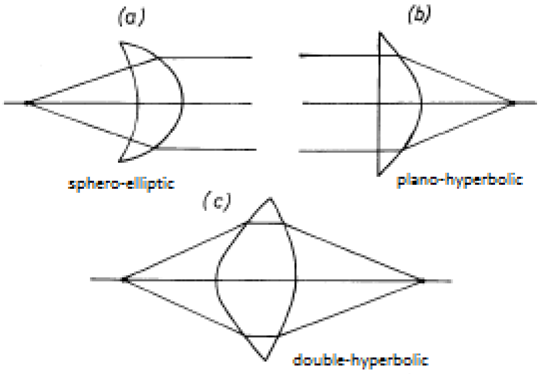


Fig.18 Three types of asferical lenses.

Unlike Cartesian surfaces and aspheric lenses, when stigmatism is achieved for a single pair of conjugate points, a ***perfect optical instrument*** (such as the Maxwellian distribution of the refractive index, called "fish-eye", see Chapter 3.3) achieves biunivocal and reciprocal correspondence between any object point P_1 in three-dimensional space and its image point P_2. Similarly, conjugate curves generate conjugate surfaces, and these, conjugate volumes. Thus, the notion of ***stigmatic image*** of extended spatial objects is introduced in geometrical optics.

Next, supposing stigmatism is achieved for a pair of points P_1, P_2, we will deduce the general condition so that it is maintained for any corresponding pair of neighboring points Q_1, Q_2 (Fig. 19). In order to do this, we will start from the definition of conjugate points, equation (69), according to which the optical path along any ray P_1PP_2 is equal to the constant $[P_1P_2]$, and that along any ray Q_1QQ_2 is equal to the constant $[Q_1Q_2]$. The condition for preservation of stigmatism for pairs of

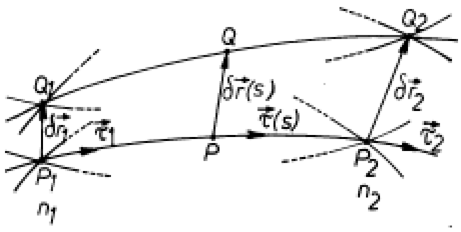


Fig. 19 Deducing the general condition for stigmatism for pairs of neighboring points.

neighboring points is thus written as

(90) $$\delta[P_1P_2] = [Q_1Q_2] - [P_1P_2] = \text{constant}.$$

But the optical path, equation (26), is the difference in phase between harmonic oscillations in the points considered, that is, $[P_1P_2] = \phi(\vec{r}_2) - \phi(\vec{r}_1)$. Because of the variation of paired conjugate points P_1, P_2, we thus obtain

$$\delta[P_1P_2] = \delta\phi(\vec{r}_2) - \delta\phi(\vec{r}_1) = \nabla\phi(\vec{r}_2)\cdot\delta\vec{r}_2 - \nabla\phi(\vec{r}_1)\cdot\delta\vec{r}_1 =$$

(91) $$= n_2\vec{\tau}_2\cdot\delta\vec{r}_2 - n_1\vec{\tau}_1\cdot\delta\vec{r}_1,$$

where we've used the eikonal equation $\nabla\phi = n\vec{\tau}$, that is, equation (9). The variations $\delta\vec{r}_1 = \overrightarrow{P_1Q_1}$ and $\delta\vec{r}_2 = \overrightarrow{P_2Q_2}$ define a new pair of neighboring points Q_1, Q_2. If (P_1, P_2) and (Q_1, Q_2) are pairs of conjugate points, then from equations (90), (91) we obtain ***the general condition for stigmatism***, or ***the cosine theorem***

(92) $$n_2\vec{\tau}_2\cdot\delta\vec{r}_2 - n_1\vec{\tau}_1\cdot\delta\vec{r}_1 = n_2\delta r_2\cos(\vec{\tau}_2,\delta\vec{r}_2) - n_1\delta r_1\cos(\vec{\tau}_1,\delta\vec{r}_1) = \text{constant}.$$

This equation links the elementary optical lengths $n_1\cdot\delta r_1$ and $n_2\cdot\delta r_2$ of the object and its stigmatic image to their orientation in the corresponding conjugate points P_1, P_2 relative to any light ray P_1PP_2 that passes through them.

The fundamental stigmatism theorem (92) can also be demonstrated if we consider the Q_1QQ_2 rays as variations of the P_1PP_2 rays (Fig. 19), so that equation (90) becomes

(93) $$\delta[P_1P_2] = \delta\int_{P_1}^{P_2} n\,ds = \int_{P_1}^{P_2}(\delta n)ds + \int_{P_1}^{P_2} n\delta(ds) = \text{constant}.$$

If we elaborate the variational calculus, as in the case of equation (28), in the end we obtain

(94) $$n_2\vec{\tau}_2\cdot\delta\vec{r}_2 - n_1\vec{\tau}_1\cdot\delta\vec{r}_1 + \int_{P_1}^{P_2}\left[\nabla n - \frac{d}{ds}(n\vec{\tau})\right]\cdot\delta\vec{r}(s)ds = \text{constant} \qquad ,$$

where the integral is eliminated thanks to ray equation (13).

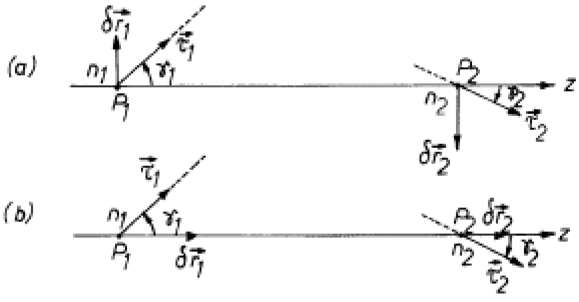


Fig. 20 Deducing the condition for transverse (a) and longitudinal (b) stigmatism.

The majority of image forming optical instruments have rotational symmetry. For this reason, we will now analyze the condition for stigmatism in the vicinity of a given arbitrary pair of conjugate points P_1, P_2, situated along the optical axis Oz of a rotating system.

Let us first consider the condition for transversal stigmatism (aplanatism) in the case of small objects and planar images, perpendicular to the optical axis (Fig. 20.a). This condition is most frequently imposed on optical instruments, and especially on microscope objectives and projection apparatuses. In this case, the general condition (92) assumes the special case form

(95) $$n_2 \delta r_2 \sin\gamma_2 - n_1 \delta r_1 \sin\gamma_1 = \text{constant}.$$

In order to determine the constant, we will use the light ray propagating along the optical axis $(\gamma_1 = \gamma_2 = 0)$, so that in the end we obtain the ***condition for transversal*** (***aplanatic***) ***stigmatism***, also called ***Abbe's sine condition***

(96) $$n_1 \delta r_1 \sin\gamma_1 = n_2 \delta r_2 \sin\gamma_2.$$

This equation must be satisfied for any ray P_1PP_2 passing through the conjugate points P_1, P_2, that is, for any pair of angles γ_1, γ_2. In the case of paraxial rays, that is, rays with small inclination γ relative to the optical axis, so that $\sin\gamma \approx \gamma$, condition (96) is reduced to the ***Lagrange-Helmholtz theorem***

(97) $$n_1 \delta r_1 \gamma_1 = n_2 \delta r_2 \gamma_2.$$

An important requirement imposed on optical systems is that the image should be similar to the object (***the property of orthoscopy***). In these conditions, the ***transversal linear magnification*** $m_t \overset{\text{def}}{=} \delta r_2 / \delta r_1$ of the system must be constant, and Abbe's sine condition is written as

(98) $$\frac{\sin\gamma_1}{\sin\gamma_2} = m_t \frac{n_2}{n_1} = \text{constant} \quad .$$

Let us next consider the condition for axial stigmatism in the case of small objects and linear images situated along the optical axis (Fig. 20.b). This condition is important in building instruments designed to form depth images or to act upon a moving point along the optical axis. In this case, the general condition (92) becomes

(99) $$n_2 \delta r_2 \cos\gamma_2 - n_1 \delta r_1 \cos\gamma_1 = \text{constant} \quad ,$$

or, if determining the constant with the help of the axial ray ($\gamma_1 = \gamma_2 = 0$),

(100) $$n_1 \delta r_1 (1 - \cos\gamma_1) = n_2 \delta r_2 (1 - \cos\gamma_2).$$

We have thus obtained the ***condition for axial stigmatism***, or ***Herschel's sine condition***

(101) $$n_1 \delta r_1 \sin^2(\gamma_1/2) = n_2 \delta r_2 \sin^2(\gamma_2/2).$$

This equation must be satisfied for any ray P_1PP_2, that is, for any pair of angles γ_1, γ_2. If we consider the ***axial linear magnification*** $m_a \overset{\text{def}}{=} \delta r_2/\delta r_1$ of the system to be constant, Herschel's sines condition is written as

(102) $$\frac{\sin^2(\gamma_1/2)}{\sin^2(\gamma_2/2)} = m_a \cdot \frac{n_2}{n_1} = \text{constant} \quad .$$

Unfortunately, the Abbe (equation (98)) and Herschel (equation (102)), conditions are not compatible with high inclination, except in the special case in which $|\gamma_1| = |\gamma_2|$ and $|m_t| = |m_a| = n_1/n_2$. In conclusion, with the exception mentioned, it is impossible to build an axial optical instrument that can form the stigmatic image of an element with volume located on the optical axis through light beams of large spread. For this reason, when building optical instruments in practice, the condition that best meets the purpose is used.

Generally, the two conditions for stigmatism (98) and (102) can only be simultaneously satisfied if the image is formed using paraxial rays ($\sin\gamma \approx \gamma$), so that

(103) $$\frac{\sin\gamma_2}{\sin\gamma_1} \approx \frac{\sin(\gamma_2/2)}{\sin(\gamma_1/2)} \approx \frac{\gamma_2}{\gamma_1} = m_u \quad ,$$

where $m_a \overset{\text{def}}{=} \gamma_2/\gamma_1$ represents ***angular magnification***. For paraxial rays there are simple equations linking angular and linear magnification. Thus, the Abbe and Herschel conditions become



(104) $$m_t m_u = n_1/n_2, \qquad m_a m_u^2 = n_1/n_2,$$

and the equation linking the three parameters is written as

(105) $$m_a m_u = m_t.$$

To conclude this chapter, we will deduce a fundamental link between the brightness of a small planar, transverse object, of area dS_1, and that of its aplanatic image, of area dS_2 (Fig. 21). According to Abbe's sine condition, equation (96), we have

(106) $$n_1^2 dS_1 \sin^2\gamma_1 = n_2^2 dS_2 \sin^2\gamma_2.$$

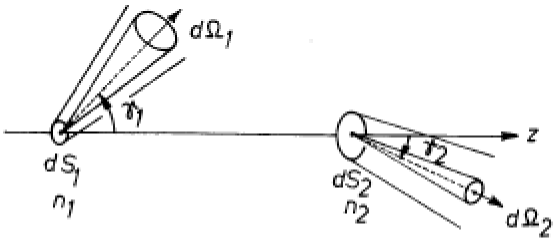


Fig. 21 Deducing the link between image brightness and object brightness.

The energetic brightness of spatially extended light sources, in any given direction γ relative to the normal to their surface, is characterized by the ***radiance*** (***brightness***, ***luminance***) $L(\gamma)$, defined as the flux of energy emitted into a solid angle unit by an apparent surface unit, that is,

(107) $$L(\gamma) = \frac{dF}{d\Omega dS \cos\gamma}.$$

In the international system, brightness is thus measured in $Watt/steradian.m^2$. The sources that satisfy ***Lambert's law*** emit completely chaotic ("randomized") light, so that their radiance does not depend on λ (black body sources of perfectly diffusing sources). In order to retain general conditions, we will keep this dependency and use the equation for conjugate fluxes written as

(108) $$dF_2(\gamma_2) = T(\gamma_1) dF_1(\gamma_1),$$

where the transmission factor $T(\leq 1)$ is determined by energy losses suffered within the system, caused by the reflection, absorption, and diffusion of light. If we introduce radiance, equation (107), and consider that $d\Omega = 2\pi \sin\gamma \, d\gamma$, we can also write equation (108) as

(109) $$T(\gamma_1) L_1(\gamma_1) dS_1 \sin\gamma_1 \cos\gamma_1 d\gamma_1 = L_2(\gamma_2) dS_2 \sin\gamma_2 \cos\gamma_2 d\gamma_2.$$

On the other hand, if we differentiate equation (106), we have

(110) $$n_1^2 dS_1 \sin\gamma_1 \cos\gamma_1 d\gamma_1 = n_2^2 dS_2 \sin\gamma_2 \cos\gamma_2 d\gamma_2.$$

Finally, from equations (109) and (110), we obtain the ***Clausius theorem***

(111) $$T(\gamma_1)\frac{L_1(\gamma_1)}{n_1^2} = \frac{L_2(\gamma_2)}{n_2^2},$$

or, if we use the expression for radiance, equation (107), and equation (108),

(112) $$n_1^2 d\Omega_1 dS_1 \cos\gamma_1 = n_2^2 d\Omega_2 dS_2 \cos\gamma_2.$$

The parameter $n^2 d\Omega dS \cos\gamma$ is called ***beam extension***. The Clausius theorem (112) thus states that ***the beam extension is conserved***. In other words, the larger the solid angle $d\Omega$, the smaller the apparent surface $dS \cos\gamma$ (see Fig. 21). This law of conservation has numerous practical consequences in photometry. For example, in the best conditions ($T = 1$), from equation (111) we obtain $L_1/n_1^2 = L_2/n_2^2$. In the special case in which $n_1 = n_2$, it follows that radiance is conserved, $L_1 = L_2$. In these conditions, not even the best focalization can increase the image's brightness past the brightness of the object. In other words, the optical system does not allow energy to pass from an apparent temperature T_1 to an apparent temperature $T_2 > T_1$.



Chapter II
CENTERED OPTICAL SYSTEMS

The most important image forming optical systems and parts (lenses, mirrors), are the ***centered optical systems***. These are a succession of homogenous, isotropic, and transparent optical media, limited by spherical surfaces with their vertex and center of curvature located on the same line. This is the system's symmetry axis, and is termed ***principal optical axis***. In practice, centered optical systems contain a large number of diopters, so that they partially compensate for chromatic aberrations (see Chapter 2.7), as well as geometric aberrations, which appear in the extra-paraxial field (see Chapter 2.8). Designing these systems entails ray tracing, by repeatedly applying the law of refraction or reflection at every separating surface and the rectilinear propagation across the homogenous media between these surfaces. This task, which is simple in principle, becomes formidable if high precision is required. That is why designing optical systems, from simple ray tracing, to correcting bothersome aberrations to meet the requirements of the intended application, is today done with the help of high-speed computers. Based on the parameters given in instructions, the computer can select the number of diopters, curvatures, refractive indices (types of optical glasses), thicknesses, apertures, and, last but not least, the weight and the price of the system.

2.1 Spherical Diopters

The easiest to fashion high-precision surface for lenses and mirrors is the spherical surface. This is why we will now study the conditions under which forming optical images is possible with the use of a ***spherical diopter***, that is, a system of two homogenous, isotropic, and transparent media separated by a spherical surface. The spherical mirror can be treated in a similar fashion.*

For simplicity, we will study the paths of light rays in a meridional plane yOz, and ***orient the optical axis Oz in the general direction of propagation of the rays of incidence, from left to right***. Let us thus consider a narrow light beam, bordered by the infinitely neighboring rays P_1PP_2'' and P_1QP_2''' emitted by the object point P_1

* ***Dioptrics*** (from Greek διά = through) is the optics of refringent systems, and ***catoptrics*** (from Greek κατα = by), the optics of reflectant systems.

situated on the optical axis; the corresponding refracted rays intersect each other in the extra-axial point P_2', and the optical axis in points P_2'' and P_2''' (see Fig. 22).

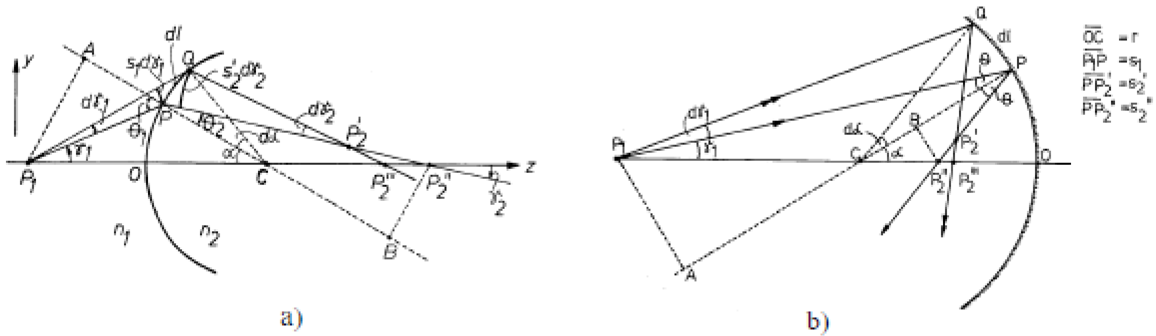


Fig. 22 Formation of images in the spherical diopter (a) and the spherical mirror (b).

We will use the notation $OC = r$ for the diopter radius, and $P_1P = s_1$, $PP_2' = s_2'$, $PP_2'' = s_2''$. If we differentiate the law of refraction $n_1 \sin\theta_1 = n_2 \sin\theta_2$ with respect to the circular arc $\widehat{\mathrm{OP}} = l$, we get

(113) $$n_1 \cos\theta_1 \frac{d\theta_1}{dl} = n_2 \cos\theta_2 \frac{d\theta_2}{dl} \quad ,$$

or, seeing that $\theta_1 = \alpha + \gamma_1$ and $\theta_2 = \alpha - \gamma_2$,

(114) $$n_1 \cos\theta_1 \left(\frac{d\alpha}{dl} + \frac{d\gamma_1}{dl}\right) = n_2 \cos\theta_2 \left(\frac{d\alpha}{dl} - \frac{d\gamma_2}{dl}\right).$$

The parameters $d\alpha/dl$, $d\gamma_1/dl$, $d\gamma_2/dl$ in the previous equation can be replaced with their expressions resulting from the equations $dl = r d\alpha$, $s_1 d\gamma_1 = dl \cos\theta_1$, $s_2' d\gamma_2 = dl \cos\theta_2$, so that, in the end, we obtain ***Young's first equation***

(115) $$\frac{n_1 \cos^2\theta_1}{s_1} + \frac{n_2 \cos^2\theta_2}{s_2'} = \frac{n_2 \cos\theta_2 - n_1 \cos\theta_1}{r} \quad ,$$

which determines the skew abscissa s_2'.

Let us next consider the similar triangles P_1AC and $P_2''BC$, obtained by drawing the perpendiculars on line CP from points P_1 and P_2''. We have $AC/BC = P_1A/P_2''B$, or

(116) $$\frac{s_1 \cos\theta_1 + r}{s_2'' \cos\theta_2 - r} = \frac{s_1 \sin\theta_1}{s_2'' \sin\theta_2} = \frac{s_1 n_2}{s_2'' n_1} \quad ,$$

from which follows ***Young's second equation***

(117) $$\frac{n_1}{s_1}+\frac{n_2}{s''_2}=\frac{n_2\cos\theta_2-n_1\cos\theta_1}{r}\quad,$$

which determines the skew abscissa s''_2.

A similar analysis to that performed above for spherical diopters leads to ***Young's equations for the spherical mirror*** (see Fig. 22.b)

(118) $$-\frac{1}{s_1}+\frac{1}{s'_2}=\frac{2}{r\cos\theta_1}\quad,$$

(119) $$-\frac{1}{s_1}+\frac{1}{s''_2}=\frac{2\cos\theta_1}{r}\quad.$$

Note that we can specialize the spherical diopter formulae (115), (117) for reflection, by simply making the formal replacements $\theta_2=\theta_1$ and $n_2=-n_1$.

Young's equations were deduced for the special case in which the object point P_1 is located on the left of the diopter (mirror) surface, while point P_2 and the center of curvature C are located on the right, as in Fig. 22.a. From here onwards, it will be convenient to adopt a ***rule for segment signs*** that will allow us to represent all possible cases using the same equations. From studying the various special cases, we can conclude that such a rule exists, and that it consists of attributing signs to the abscissas of points P_1, P_2, C, so that $s_1>0$, $s_2<0$, $r<0$ if these points are located to the left of the diopter (mirror) surface, and so that we have the opposite situation if they are located to its right. Based on this convention, in the case of diopters, points P_1, P_2 are ***real*** if $s_1, s_2>0$, and ***virtual*** if $s_1, s_2<0$, and in the case of mirrors, they are ***real*** if $s_1>0$, $s_2<0$, and ***virtual*** if $s_1<0$, $s_2>0$.

The difference $\delta=s''_2-s'_2$, which describes the deviation from stigmatism, is called ***astigmatic distance***. Usually, spherical diopters and mirrors aren't rigorously stigmatic, meaning $\delta\neq0$ ($s'_2\neq s''_2$). However, there indeed are exceptional cases of ***rigorous stigmatism***, such is that of the ***Weierstrass points*** (***Young points***). In these cases, from equations (115), (117), and the condition for stigmatism $s'_2=s''_2=s_2$, it follows that

(120) $$s_1=\left(1-\frac{n_2^2}{n_1^2}\right)\frac{n_1 r}{n_2\cos\theta_2-n_1\cos\theta_1}\quad,$$

(121) $$s_2=\left(1-\frac{n_1^2}{n_2^2}\right)\frac{n_2 r}{n_2\cos\theta_2-n_1\cos\theta_1}\quad.$$

In this case, points P'_2, P''_2, P'''_2 are the same point P_2, located on the optical axis.

As for the axial ray ($\theta_1 = \theta_2 = 0$), the skew abscissas become regular abscissas $s_1 = P_1O = p_1$, $s_2 = OP_2 = p_2$, where

(122)
$$\begin{cases} p_1 = -\left(1 + \frac{n_2}{n_1}\right) r, \\ p_2 = \left(1 + \frac{n_1}{n_2}\right) r, \end{cases}$$

determine ***the position of conjugate Weierstrass points***. In Fig. 23 we have an illustration of a case in which $n_1 = 2$, $n_2 = 1$, and, therefore, $p_1 = -3r/2$, $p_2 = 3r$, where $r < 0$. From equation (122), it follows that $r > 0$ implies that $p_1 < 0$, $p_2 > 0$, and $r < 0$ implies that $p_1 > 0$, $p_2 < 0$, meaning ***the Weierstrass points are located on the same side as the center of curvature C***.

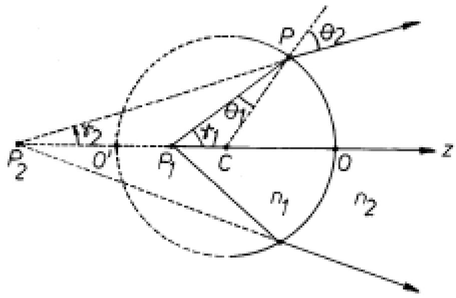


Fig. 23 The conjugate (Weierstrass, Young) points of a spherical diopter.

Expressions (122) can also be written as $CP_1 = (n_2/n_1)r$, $CP_2 = (n_1/n_2)r$, from which it follows that $CP_1 \cdot CP_2 = r^2$, meaning points P_1, P_2 are harmonically conjugate with points O, O',
in which the diopter surface intersects the optical axis. We can also write it as $CP_1/CP = CP/CP_2$, meaning triangles CP_1P and CPP_2 are similar, $\gamma_1 = \theta_2$, and $\gamma_2 = \theta_1$, therefore,

$$\frac{\sin\gamma_1}{\sin\gamma_2} = \frac{\sin\theta_2}{\sin\theta_1} = \frac{n_1}{n_2} = \text{constant} \quad .$$

We've thus demonstrated that the Weierstrass points satisfy Abbe's sine condition (equation (98))

$$\frac{\sin\gamma_1}{\sin\gamma_2} = m_t \frac{n_2}{n_1} = \text{constant} \quad ,$$

with transversal linear magnification $m_t = (n_1/n_2)^2$.

Contemporary
Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017

The aplanatism property of Weierstrass points is used in fashioning microscope objectives of high ***numerical aperture*** $n_1 \sin\gamma_1$. To this end, the cone of light collected from every object point P_1 must be as wide as possible, as is illustrated in Fig. 24.a,b. Fig.24.a represents ***the Amici method***, which uses convex-concave lenses in which object point P_1 is simultaneously located both at the center of spherical diopter $1'$ and at the first Weierstrass point of diopter 1. The second Weierstrass point P_2 of diopter 1 thus represents the aplanatic image of object point P_1 in the first lens $(1, 1')$. Point P_3 similarly represents the aplanatic image of point P_2 in the second lens $(2, 2')$. Thus, satisfying the condition of rigorous aplanatism, the wide light beam with its apex at object point P_1, with an aperture nearing the theoretical value $2\gamma_1 = \pi$, is transformed in a beam of paraxial rays, with its apex at point P_3. The objective usually ends with an achromatic system (A), that corrects chromatic aberrations (see Chapter 2.7). Further increase of the numerical aperture $n_1 \sin\gamma_1$ is achieved by introducing between the object and frontal lens an immersion liquid of high refractive index n_1, usually cedar oil ($n = 1.515$), which basically has the same refractive index as that of glass. In this way, losses due to reflection at the first surface of the frontal lens are also reduced. Fig. 24.b illustrates such an ***immersion objective***. In this case, the first diopter $1'$ no longer plays a role, so that the frontal lens can be a planar-convex lens. We've insisted on this issue because it represented an important moment in the evolution of image forming optical instruments. Within the scalar theory of light diffraction, it is demonstrated that the minimum distance $(\delta r_1)_{min}$ between two points on the object that can still be resolved is limited by the phenomenon of diffraction. This minimum distance is given by the ***Abbe formula***

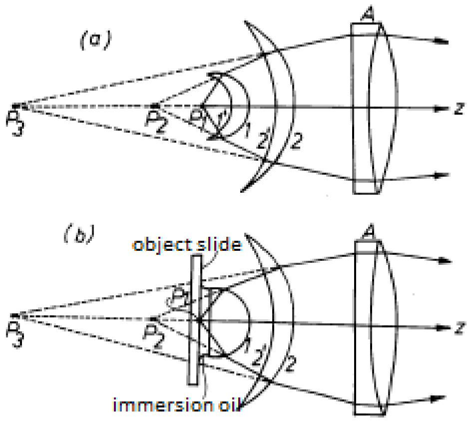


Fig. 24 Microscope objectives with high numerical aperture: a) of type Amici, B) of an immersion type.

(123) $$(\delta r_1)_{\min} = \frac{0,61\lambda_1}{\sin\gamma_1} = \frac{0,61\lambda_0}{n_1 \sin\gamma_1}.$$

It follows that the ***spatial resolving power of microscope objectives***, defined as $1/(\delta r_1)_{min}$, can be increased by using radiation of smaller wavelength and attaining a higher numerical aperture $n_1 \sin\gamma_1$.

We've considered the exceptional case of rigorous stigmatism of Weierstrass points with wide beams. In the case of paraxial rays ($\sin\gamma \approx \gamma$), as we've seen in

Chapter 1.3, the Abbe and Herschel conditions can always be approximately satisfied, which thus allows achieving an ***approximate stigmatism*** for all points on and around the optical axis. In the particular case of spherical diopters, the paraxial approximation* gives us $s_1 \approx p_1$, $s_2' \approx s_2'' \approx p_2$, $\cos\theta \approx 1$, so that Young's equations (115), (117) lead to a single equation,

(124) $$\frac{n_1}{p_1} + \frac{n_2}{p_2} = \frac{n_2 - n_1}{r} \quad ,$$

which is the ***conjugate point relationship***. The right-hand member of this equation depends only on the system parameters (n_1, n_2, r). In the case of spherical mirrors $(n_2 = -n_1)$, relation (124) becomes

(125) $$-\frac{1}{p_1} + \frac{1}{p_2} = \frac{2}{r} \quad .$$

The conjugate point relationship (124) allows us to determine for any spherical diopter or mirror $(n_2 = -n_1)$ with a convex $(r > 0)$, concave $(r < 0)$, or planar $(r \to \infty)$ surface the positions of conjugate points p_1, p_2 relative to the surface vertex O, as well as the real or virtual character of those points, using the rule adopted above for assigning plus or minus signs.

If in the conjugate point relationship (124) $p_2 \to \infty$, then $p_1 \to f_1$, and if $p_1 \to \infty$, then $p_2 \to f_2$, where

(126) $$f_1 = \frac{n_1 r}{n_2 - n_1}, \quad f_2 = \frac{n_2 r}{n_2 - n_1}$$

are termed ***focal distances*** (object and image, respectively). Obviously, it follows that

(127) $$f_2 - f_1 = r, \quad f_1/f_2 = n_1/n_2.$$

The focal distances determine the corresponding ***focal points***, or ***foci***, F_1, F_2. These are ***real*** if $f > 0$, or ***virtual*** if $f < 0$.

The inverses of the focal distances, that is, $C_1 = 1/f_1$ and $C_2 = 1/f_2$, are called ***convergences***. The diopter is convergent or divergent according to whether its

* The first order ***paraxial approximation***, or Gaussian paraxial approximation (named after Gauss, who first used it systematically in ***Dioptrische Untersuchungen***, (1843)) is applicable when the light beam rays and the normals at the refringent surfaces form small (but otherwise arbitrary) angles x with the optical axis, so that $\sin x \simeq \tan x \simeq x$, and $\cos x \simeq 1$, which practically means that $x \leq 0.1\ radians \simeq 6°$.

convergence is positive of negative. The usual unit of convergences is the ***diopter*** (m^{-1}).

An equivalent form of the conjugate point relationship (124) can be obtained by dividing with $(n_2 - n_1)/r$ and introducing focal distances. We get

(128) $$\frac{f_1}{p_1} + \frac{f_2}{p_2} = 1 \quad ,$$

an equation known as the ***Huygens-Gauss formula***.

Finally, we can obtain another form of the conjugate point relationship if we determine the positions of conjugate points P_1, P_2 using segments ζ_1, ζ_2, delimited by the corresponding focal points. By transforming the coordinates (see Fig. 25),

(129) $$\begin{cases} p_1 = \zeta_1 + f_1, \\ p_2 = \zeta_2 + f_2, \end{cases}$$

from the Huygens-Gauss formula (128), the following simple and symmetrical relation immediately follows:

(13) $$\zeta_1 \zeta_2 = f_1 f_2 \quad ,$$

an equation which is called the ***Newton formula***.

The equations deduced above are also applicable for the spherical mirror ($n_2 = -n_1$), so for focal distances

(131) $$f_2 = -f_1 = \frac{r}{2} \quad .$$

In this case, foci F_1, F_2 are situated on the same side of the mirror surface and coincide in a common focal point, located at half the distance between the mirror vertex and the center of curvature. This single focus is real for concave, and virtual for convex mirrors.

Up to this point we've considered the relationship between conjugate points P_1, P_2 situated along the principal optical axis (which passes through the center of curvature C and the vertex O of the spherical cap). Obviously, this relationship also exists between conjugate points Q_1, Q_2, located on any other secondary axis (which passes through C, but not through O), as is illustrated in Fig. 25. For this reason, the image of a circle arc of spherical cap $\widehat{P_1Q_1}$ is another arc or cap $\widehat{P_2Q_2}$, both having their center located at C. However, in paraxial approximation, we only consider the object and image points in the vicinity of the principal optical axis, so that the arcs and caps

$\widehat{PQ}$ are confounded with the small transverse objects or images δr, to which they are tangent in the corresponding points P.

The images are graphically constructed in a very convenient way, using rays of construction (***principal rays***), which pass through the foci and the center of curvature (Fig. 25). This construction is consistent with the conjugate point relationship. Thus, from the similarity between triangles $P_1Q_1F_1$ and OR_2F_1 or $P_2Q_2F_2$ and OR_1F_2, it follows that

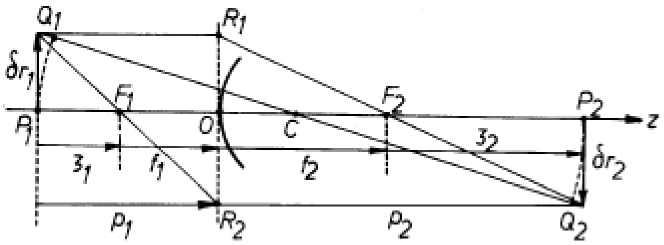


Fig. 25 Forming the image of an object in a spherical diopter.

(132)
$$m_t \overset{def}{=} \frac{\delta r_2}{\delta r_1} = -\frac{f_1}{\zeta_1} = -\frac{\zeta_2}{f_2} \quad ,$$

where the last equality confirms the relation $\zeta_1\zeta_2 = f_1 f_2$ (equation (130)). Another expression for the ***transversal linear magnification*** results from the similarity between triangles P_1Q_1C and P_2Q_2C, based on which we obtain the equation $\delta r_2/\delta r_1 = -(p_2 - r)/(p_1 + r)$, or, if we eliminate r using equation (124)

(133)
$$m_t \overset{def}{=} \frac{\delta r_2}{\delta r_1} = -\frac{n_1}{n_2} \cdot \frac{p_2}{p_1}.$$

Note that the transversal magnification m_t can be very high if $p_2$is very large, that is, if P_1 is near focus F_1.

The conjugate point relationship (124) and expression (133) allow us to determine the position and size of image p_2, δr_2, based on the position and size of object p_1, δr_1, through the following equations:

(134)
$$\begin{cases} p_2 = \frac{n_2 r p_1}{p_1(n_2-n_1)-rn_1}, \\ \delta r_2 = \frac{n_1 r \delta r_1}{p_1(n_2-n_1)-rn_1}. \end{cases}$$

Successively applying these formulae for each diopter is a direct method of constructing images in centered optical systems in paraxial approximation. Thus, the image formed by the first surface acts as an object for the second surface, the image formed by the second surface acts as an object for the third, and so on.

The ***axial linear magnification*** results from differentiating the conjugate point

relationship (124), and is written as:

(135) $$m_a \stackrel{def}{=} -\frac{dp_2}{dp_1} = \frac{n_1}{n_2} \cdot \frac{p_2^2}{p_1^2} \quad .$$

Finally, if we consider an arbitrary paraxial ray P_1PP_2 extending between two conjugate points located on the principal optical axis (see Fig. 22), we have $\widehat{OP} = p_1\gamma_1 = -p_2\gamma_2$, from which we obtain the ***angular magnification***

(136) $$m_u \stackrel{def}{=} \frac{\gamma_2}{\gamma_1} = -\frac{p_1}{p_2} \quad .$$

By multiplying this equation by equation (133), we obtain our confirmation of the general relationship $m_1 m_u = n_1/n_2$ (equation (104)), and of the Lagrange-Helmholtz theorem $n_1\delta r_1\gamma_1 = n_2\delta r_2\gamma_2$ (equation 97). In fact, the three spherical diopter magnifications also serve to confirm the other general relationships, $m_a m_u^2 = n_1/n_2$ and $m_a m_u = m_t$, applicable to any axial system in the paraxial approximation (also see Chapter 1.3).

2.2 The Transfer Matrix

Next, we will elaborate the paraxial analysis of centered optical systems. This issue has significant practical importance, since the images obtained through paraxial rays have no geometric aberrations, and the formulae deduced in this approximation are sufficiently precise for numerous applications. On the other hand, they constitute the foundation stone for more exact calculations, because deviations from the paraxial analysis formulae are a convenient measuring tool for appreciating the quality of real optical instruments.

As we've seen above, one direct method of achieving a paraxial analysis of centered optical systems involves successively applying the (134) relations separately for each diopter. Alternatively, because of the repetitive way in which this type of linear relations describing rectilinear propagation and refraction (or reflection) occur, ***any optical system may be associated with a transfer matrix***, which is calculated as a simple product between fundamental translation and refraction (or reflection) matrices. The matrix method is a powerful instrument for calculating and designing optical systems, and, as we will soon see, it allows the demonstration of important theorems of paraxial geometric optics.

For convenience, we will now consider the meridional paraxial rays in plane

yOz with the Cartesian axes to be oriented according to the segment sign rule described in the spherical diopter chapter (see Chapter 2.1). We will thus retain the convention according to which axis Oz is situated along the principal optical axis and is oriented in the general direction of propagation of the rays of incidence, considered to be from left to right, and the Oy axis to be oriented from downwards upwards. We will also adopt the convention of measuring the inclination γ of light rays in radians relative to orientation of axis Oz, and of assigning a sign to it according to trigonometric direction.

It will be convenient to define the ***state of the light ray*** in any one of its points using the 2×1 column matrix (the ***state vector)***

(137) $$V = \begin{bmatrix} y \\ \Gamma \end{bmatrix} \quad ,$$

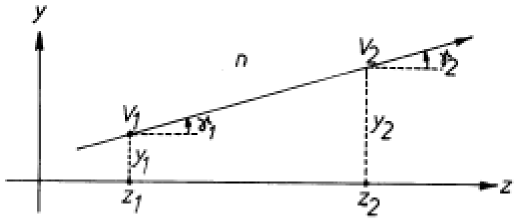


Fig. 26 Matrix description of rectilinear propagation.

where y is the ***distance to the optical axis***, and $\Gamma = n\gamma$ is the ***reduced inclination***.

The rectilinear propagation, or translation, between planes $z = z_1$ and $z = z_2$ is described by the following equations (see Fig. 26):

(138) $$\begin{cases} y_2 = y_1 + (z_2 - z_1) tg\gamma_1, \\ \gamma_2 = \gamma_1. \end{cases}$$

By introducing the ***translation*** $t = z_2 - z_1$, the ***reduced translation*** t/n, the ***reduced inclinations*** $\Gamma_1 = n\gamma_1$, $\Gamma_2 = n\gamma_2$, and taking into account the paraxial approximation $\tan\gamma \approx \gamma$, we can write the translation transformations (138) as

(139) $$\begin{cases} y_2 = 1 \cdot y_1 + (t/n)\Gamma_1 \\ \Gamma_2 = 0 \cdot y_1 + 1 \cdot \Gamma_1 \end{cases} \quad ,$$

or, in a shorter form, $V_2 = TV_1$, where the square 2×2 matrix

(140) $$T = \begin{bmatrix} 1 & t/n \\ 0 & 1 \end{bmatrix} \quad ,$$

is the ***translation matrix***. It contains all the information regarding the system traversed by the ray, that is, the optical medium of thickness t and refractive index n. If the translation is null ($t = 0$), the translation matrix becomes the unit matrix, and $\gamma_2 = \gamma_1$, $\Gamma_1 = \Gamma_2$. Note that the translation matrix is a unitary matrix, meaning $\det(T) =$



1.

Let us next determine the matrix operator which represents refraction, that is, the linear relationship between state vector $V_2(y_2, n_2\gamma_2)$ at the beginning of the refracted ray, and the state vector $V_1(y_1, n_1\gamma_1)$ at the end of the ray of incidence. The first relationship is simply

(141) $$y_2 = y_1 \quad .$$

The second linear relation follows from the law of refraction in paraxial approximation $n_1\theta_1 = n_2\theta_2$, that is, $n_1(\alpha + \gamma_1) = n_2(\alpha + \gamma_2)$, or:

(142) $$n_1\left(\frac{y_1}{r} + \gamma_1\right) = n_2\left(\frac{y_1}{r} + \gamma_2\right) \quad ,$$

as can be seen in Fig. 27. From here we deduce the desired relationship

(143) $$n_2\gamma_2 = \frac{n_1 - n_2}{r} y_1 + n_1\gamma_1 \quad .$$

Note that this equation is actually the conjugate point relationship for spherical diopters, that is, equation (124), since $\gamma_1 = y_1/p_1$ and $\gamma_2 = -y_1/p_2$. The refraction transformations (141), (143) can be written as

(144) $$\begin{cases} y_2 = 1 \cdot y_1 + 0 \cdot \Gamma_1 \\ \Gamma_2 = Q \cdot y_1 + 1 \cdot \Gamma_1 \end{cases} \quad ,$$

or, in a shorter form, $V_2 = RV_1$, where the 2×2 square matrix

(145) $$R = \begin{bmatrix} 1 & 0 \\ Q & 1 \end{bmatrix} \quad , \quad Q = \frac{n_1 - n_2}{r} \quad ,$$

is the ***refraction matrix***. The parameter Q, which contains all the information regarding the diopter, that is, n_1, n_2, r, is termed ***refraction power***. If $Q = 0$ (if $n_1 = n_2$ or $r \to \infty$), the refraction matrix becomes a unit matrix, so that $y_1 = y_2$, $\Gamma_1 = \Gamma_2$. Note that it's unitary matrix, meaning $\det(R) = 1$. The relations of the spherical diopter are also applicable in the case of the spherical mirror ($n_2 = -n_1$), in which the parameter $Q = 2n_1/r$ is, however, called ***reflection power***.

Let us now consider the general case of a paraxial ray passing through a centered optical system formed out of m spherical surfaces separated by $m - 1$ media of different refractivity (Fig. 28). Such an axial dioptric system is considered to extend between the entrance plane, tangent to the first refractive surface Σ_1 at its vertex, and

the exit plane, tangent to the last refractive surface Σ_m at its vertex. As the light ray progresses along the system, we have

$$\begin{aligned} V_1'' &= R_1 V_1' \quad , \\ V_2' &= T_1 V_1'' = T_1 R_1 V_1' \quad , \\ V_2'' &= R_2 V_2' = R_2 T_1 R_1 V_1' \quad , \\ &\dots\dots\dots\dots\dots\dots \\ V_m'' &= R_m V_m' = R_m T_{m-1} R_{m-1} \dots R_2 T_1 R_1 V_1' \quad . \end{aligned}$$

The last equation links the exit vector V_m'' to the entrance vector V_1' through the ***transfer equation***

(146) $$V_m'' = S V_1' \quad ,$$

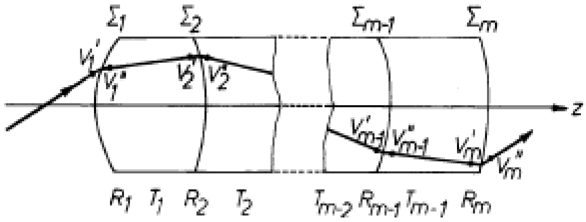


Fig. 28 A paraxial ray passing through a centered optical system.

where

(147) $$S = R_m T_{m-1} R_{m-1} \dots R_2 T_1 R_1$$

is the ***transfer matrix of the system***, considered to extend between the vertices (planes) of entrance and exit. Note that the matrix S associated with the system is the product of individual refraction and translation matrices calculated in ***inverse order***, meaning in an order opposite that in which light propagates throughout the system. Since the refraction and translation matrices are 2×2 square unitary matrices, it follows that the product matrix is of the same type, that is, its form is

(148) $$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \quad ,$$

where

(149) $$\det S = S_{11} S_{22} - S_{12} S_{21} = 1 \quad .$$

The elements of matrix S contain all the information regarding the optical system's parameters, namely the refractive indices, the curvature radii, and the thicknesses of the diopter components. Due to relation (149), it follows that only three of the four elements of matrix S are independent, and, as we will see in the following chapters, they determine all the properties of the system as an image-forming



instrument. Property (149) is an important tool for verifying the correctness of calculations, during the calculating process and at its end. Thus, as we calculate the product of a long series of matrices, it is advisable to check, now and then, if the determinant of the product matrix is equal to unity. If it is not, it means an error has occurred in calculations.

Due to the associative property, when effectively calculating matrix S, there are multiple methods of obtaining the product. The most convenient is to first associate the matrices in pairs. But, since in a general case we have alternative product sequences RT, our calculations must follow the inverse order discussed above, because, as can readily be verified, the product RT is not commutative, that is, $RT \neq TR$. Only in the special cases of successive translation through parallel refractive plane layers, or of successive refraction through attached diopters (thin lenses), the corresponding products are commutative, and so we have

(150) $$S = \prod_{i=1}^{m} T_i = \begin{bmatrix} 1 & \sum_{i=1}^{m} t_i/n_i \\ 0 & 1 \end{bmatrix} ,$$

and, respectively,

(151) $$S = \prod_{i=1}^{m} R_i = \begin{bmatrix} 1 & 0 \\ \sum_{i=1}^{m} Q_i & 1 \end{bmatrix} .$$

Of course, changing the order of the component refractive elements affects the trajectory of the ray within the system, but not matrix S, so that the relationship between the exit and entrance vectors remains unchanged.

To illustrate, let us calculate the transfer matrix for a biconvex glass lens ($n_2 = 1.5$), with curvature radii $r_1 = +2cm$ and $r_2 = -1cm$, of thickness $g = 0.5\ cm$, immersed in air $n_1 = n_3 = 1$ (Fig. 29.a).

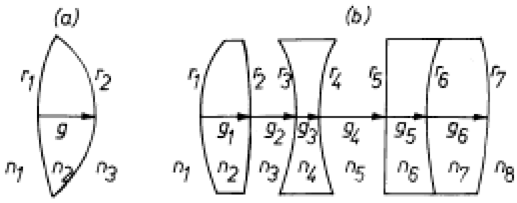


Fig. 29 Calculating the transfer matrix of a lens (a), and of a Tessar system (b).

$$S = R_2 T R_1 = \begin{bmatrix} 1 & 0 \\ Q_2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & g/n_2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ Q_1 & 1 \end{bmatrix} ,$$

where $Q_1 = (n_1 - n_2)/r_1$, $Q_2 = (n_2 - n_3)/r_2$. If we introduce the numerical data and calculate the matrix product, we obtain

(152)
$$S = \begin{bmatrix} 0,917 & 0,333 \\ -0,708 & 0,833 \end{bmatrix} .$$

Obviously, $\det(S) = 1$.

The method is similar for optical systems, no matter how complex. For example, let us consider a ***Tessar system*** (Fig. 29.b), whose data is listed in the following table:

Nr.	n	r (cm)	g (cm)
1	1	1,628	0,357
2	1,6116	-27,570	0,189
3	1	-3,457	0,081
4	1,6053	1,582	0,325
5	1	∞	0,217
6	1,5123	1,920	0,396
7	1,6116	-2,400	
8	1		

This system is free of geometric (astigmatism and field curvature) and chromatic aberrations, and many modern photographic objectives are variants. In this case, matrix S is given by the product of 13 fundamental matrices, namely

$$S = R_7 T_6 R_6 T_5 R_5 T_4 R_4 T_3 R_3 T_2 R_2 T_1 R_1$$

The calculations themselves, easily done with the help of a Fortran program, lead to the results

(153)
$$S = \begin{bmatrix} 0,867 & 1,338 \\ -0,198 & 0,848 \end{bmatrix} .$$

Obviously, we have $\det(S) = 1$.

2.3 Cardinal Elements

In order to analyze the image forming properties, it will be convenient to consider the relationship between two ray state vectors, V_1 and V_2, located within



arbitrary reference planes, $z = z_1$ and $z = z_2$, respectively, usually different from the planes tangent at the entrance and exit vertices of the optical system (see Fig. 30). The transfer equation (146) readily generalizes in the form

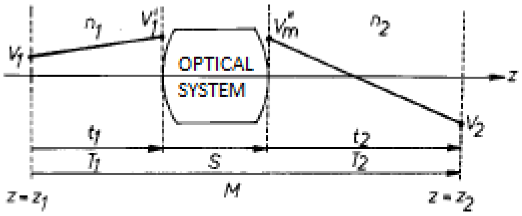


Fig. 30 Determining the transfer matrix of a centered system bordered by two random reference planes ($z = z_1$ and $z = z_2$).

(154) $$V_2 = MV_1 \quad ,$$

where the transfer matrix between the planes considered is

(155) $$M = T_2 S T_1 = \begin{bmatrix} 1 & t_2/n_2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \cdot \begin{bmatrix} 1 & t_1/n_1 \\ 0 & 1 \end{bmatrix} \quad ,$$

and has the elements

(156) $$M_{11} = S_{11} + \frac{S_{21}}{n_2} t_2 \quad ,$$

(157) $$M_{22} = S_{22} + \frac{S_{21}}{n_1} t_1 \quad ,$$

(158) $$M_{21} = S_{21} \quad ,$$

(159) $$M_{12} = \frac{S_{11}}{n_1} t_1 + \frac{S_{22}}{n_2} t_2 + \frac{S_{21}}{n_1 n_2} t_1 t_2 + S_{12} \quad ,$$

with the property that

(160) $$\det M = M_{11} M_{22} - M_{12} M_{21} = 1 \quad .$$

Note the invariance of element S_{21} with respect to the translation of the reference planes.

The transfer equation (154) represents the linear transformations

(161) $$y_2 = M_{11} y_1 + M_{12} \Gamma_1 \quad ,$$

(162) $$\Gamma_2 = M_{21} y_1 + M_{22} \Gamma_1 . \quad .$$

According to property (160), two elements of matrix M at most can be null. Let us analyze the significance of the successive elimination of these elements. When a diagonal element is eliminated, equations (160), (161), (162) become, respectively,

$$\det M = -M_{12}M_{21} = -M_{12}S_{21} = 1,\ y_2 = M_{12}\Gamma_1 \qquad (\text{if } M_{11} = 0),$$
$$\Gamma_2 = M_{21}y_1 \qquad (\text{if } M_{22} = 0),$$

from where it follows that

(163) $$y_1 = (n_2/S_{21})\gamma_2 \equiv -f_2\gamma_2 \quad (\text{if } M_{22} = 0),$$

(164) $$y_2 = -(n_1/S_{21})\gamma_1 \equiv f_1\gamma_1 \quad (\text{if } M_{11} = 0).$$

As we will see later (see equations (182), (183)), the parameters

(165) $$\begin{cases} f_1 \overset{def}{=} -n_1/S_{21}, \\ f_2 \overset{def}{=} -n_2/S_{21}, \end{cases}$$

whose values only depend on the element S_{21} of the optical system, are the object and image ***focal distances***, respectively. The geometric interpretation of equations (163), (164) can be seen in Fig. 31. It consists of a homocentric beam, transformed into a parallel beam (Fig. 31.a), or a parallel beam transformed into one homocentric (Fig. 31.b). The ***position of the focal planes*** (object and image) results from the corresponding conditions, $M_{22} = 0$ and $M_{11} = 0$, for which we use the expressions (156) and (157), namely

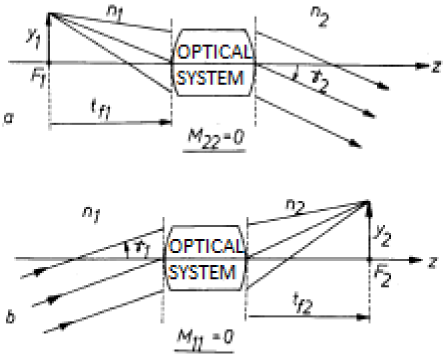


Fig. 31 The geometric interpretation of equations (163) and (164).

(166) $$\begin{cases} t_{f_1} = -(n_1/S_{21})S_{22} \equiv f_1 S_{22}, \\ t_{f_2} = -(n_2/S_{21})S_{11} \equiv f_2 S_{11}. \end{cases}$$

The intersections between the focal planes and optical axis determine the object (F_1) and image (F_2) ***focal points***, or ***foci***.

Let us next analyze the implications of eliminating element 21, that is, $M_{21} = S_{21} = 0$, so that the focal planes (equations (166)) move to infinity, and the focal distances (equations (165)) become infinite. The optical systems which have this property are termed ***afocal***, or ***telescopic***, systems. In this case, equations (156), (157) become $M_{11} = S_{11}$, $M_{22} = S_{22}$, and property (160) is written as $\det M = M_{11}M_{22} =$

$S_{11}S_{22} = 1$. Moreover, from equation (162) it follows that $\Gamma_2 = S_{22}\Gamma_1$, meaning that

(167) $$n_2\gamma_2 = S_{22}n_1\gamma_1 \quad .$$

This relation expresses the property of afocal optical systems of transforming a parallel beam of inclination γ_1 into a parallel beam of inclination γ_2 (Fig. 32).

The most important parameter describing afocal systems is the ***angular magnification***. As can be deduced from equation (167), we have

(168) $$m_u \stackrel{def}{=} \gamma_2/\gamma_1 = n_1S_{22}/n_2 \equiv n_1/n_2S_{11} \quad .$$

Note that, unlike any other optical system, in the case of telescopic systems, the angular magnification is the same for all light rays.

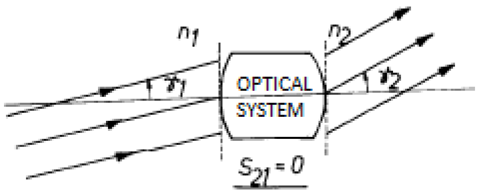


Fig. 32 Afocal (telescopic) optical system.

Finally, let us analyze the significance of condition $M_{12} = 0$. In this case, equation (160) becomes $\det M = M_{11}M_{22} = 1$, and equation (161) is written as

(169) $$y_2 = M_{11}y_1 = \frac{1}{M_{22}}y_1 \quad .$$

This means that, no matter the inclination Γ_1, a conical beam with its apex in plane $z = z_1$ is transformed into a conical beam with its apex in plane $z = z_2$ (see Fig. 33). In other words, the $M_{12} = 0$ condition is the ***stigmatism condition*** (in paraxial approximation). Referring to expression (159) of element M_{12}, we obtain the general form of the ***conjugate planes relationship***

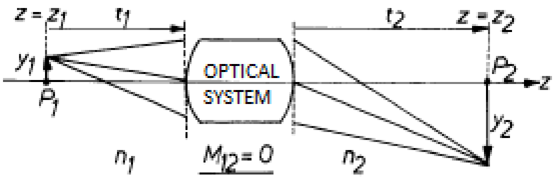


Fig. 33 The geometric interpretation of relation $M_{12} = 0$ (conjugate planes).

(170) $$\frac{n_1}{t_1}S_{22} + \frac{n_2}{t_2}S_{11} + \frac{n_1n_2}{t_1t_2}S_{12} + S_{21} = 0 \quad .$$

The conjugate planes intersect the optical axis in the conjugate object (P_1) and image (P_2) points.

Note that if $t_2 \to \infty$, then $t_1 \to t_{f_1}$, and if $t_1 \to \infty$, then $t_2 \to t_{f_2}$ (see equation 166). Therefore, each focal plane is conjugate with the corresponding plane located at

infinity. The focal planes themselves do not form a pair of conjugate planes, because t_{f_1} and t_{f_2} do not simultaneously satisfy relation (170) of conjugate planes.

Equation (169) directly gives us the ***transversal linear magnification***

(171) $$m_t \overset{def}{=} \frac{y_2}{y_1} = M_{11} = \frac{1}{M_{22}} \qquad ,$$

or, if we refer to expressions (156), (157), and to the definition of focal distances (165),

(172) $$\begin{cases} m_t = S_{11} + \frac{S_{21}}{n_2} t_2 \equiv S_{11} - \frac{t_2}{f_2}, \\ m_t^{-1} = S_{22} + \frac{S_{21}}{n_1} t_1 \equiv S_{22} - \frac{t_1}{f_1}. \end{cases}$$

Note that, since it is always true that $M_{21} = S_{21}$, equation (158), the transfer matrix between two conjugate planes has the following structure:

(173) $$M_{P_1P_2} = \begin{bmatrix} m_t & 0 \\ S_{21} & m_t^{-1} \end{bmatrix} \quad .$$

By definition, the ***principal planes*** of the optical system are that pair of conjugate planes for which the transversal linear magnification has the value $m_t = +1$, meaning that the image has the same size and the same orientation as the object. The ***positions of principal planes*** (object and image) thus follows from the conditions $M_{11} = M_{22} = 1$, in which we use expressions (156) and (157), that is

(174) $$\begin{cases} t_{p_1} = -\frac{n_1}{S_{21}}(S_{22} - 1) \equiv f_1(S_{22} - 1), \\ t_{p_2} = -\frac{n_2}{S_{21}}(S_{11} - 1) \equiv f_2(S_{11} - 1). \end{cases}$$

We may verify that positions (174) indeed satisfy the conjugate plane relationship (170), since they are, by definition, the conjugate pair for which $m_t = +1$. The intersections between the principal planes and the optical axis determine the ***principal*** object (H_1) and image (H_2) ***points***. The distance H_1H_2 between principal planes is termed ***interstitium***.

From expressions (165), (166), (174) result the following general relationships between the positions of focal and principal points, and focal distances,

(175) $$\begin{cases} t_{f_1} - t_{p_1} = f_1, \\ t_{f_2} - t_{p_2} = f_2. \end{cases}$$

Note that the transfer matrix (173) written for the pair of principal planes $(m_t = +1)$ assumes the simple form of the fundamental matrix for spherical diopters (see equation (145)), that is,

(176) $$M_{H_1H_2} = \begin{bmatrix} 1 & 0 \\ S_{21} & 1 \end{bmatrix} \quad .$$

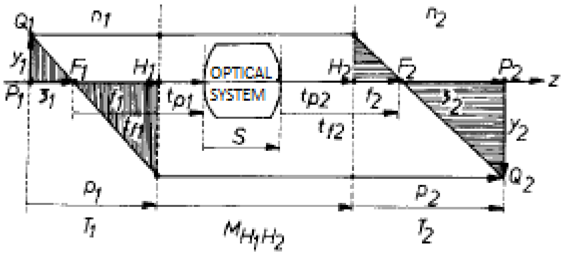


Fig. 34 The principal planes of an optical system and their usage in image construction.

Next we will use this remarkable property to simplify the formulae, by adopting the convention of measuring the object (p_1) and image (p_2) distances relative to the corresponding principal plane (Fig. 34). The transfer matrix between the conjugate planes is thus written as

(177) $$M = T_2 M_{H_1H_2} T_1 = \begin{bmatrix} 1 & p_2/n_2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ S_{21} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & p_1/n_1 \\ 0 & 1 \end{bmatrix} \quad ,$$

and has simpler elements than matrix (155), namely

(178) $$M_{11} = 1 + \frac{S_{21}}{n_2} p_2 \equiv 1 - \frac{p_2}{f_2} \quad ,$$

(179) $$M_{22} = 1 + \frac{S_{21}}{n_1} p_1 \equiv 1 - \frac{p_1}{f_1} \quad ,$$

(180) $$M_{21} = S_{21} \quad ,$$

(181) $$M_{12} = \frac{p_1}{n_1} + \frac{p_2}{n_2} + S_{21}\frac{p_1 p_2}{n_1 n_2} \equiv -\frac{1}{S_{21}}\left(\frac{p_1}{f_1} + \frac{p_2}{f_2} - \frac{p_1 p_2}{f_1 f_2}\right) \quad .$$

The stigmatism condition $M_{12} = 0$ leads to a ***generalization of the Huygens-Gauss formula***, in the form deduced for spherical diopters, equation (128), that is,

(182) $$\frac{f_1}{p_1} + \frac{f_2}{p_2} = 1 \quad ,$$

where the focal distances f_1, f_2 are considered expressed as in (165). Moreover, referring to the transformation (see formula 129))

$$\begin{cases} p_1 = \zeta_1 + f_1, \\ p_2 = \zeta_2 + f_2, \end{cases}$$

where the segments ζ_1, ζ_2 determine the positions of conjugate points P_1, P_2 relative

to the corresponding foci (see Fig. 34), from equation (182) we obtain the ***generalization of the Newton formula***

(183) $$\zeta_1 \zeta_2 = f_1 f_2 \quad .$$

The ***transversal linear magnification*** $m_t = M_{11} = 1/M_{22}$, equation (171), using elements M_{11}, M_{22} of matrix (177), can be written as

(184) $$\begin{cases} m_t = 1 + \frac{S_{21}}{n_2} p_2 \equiv 1 - \frac{p_2}{f_2} = -\frac{\zeta_2}{f_2}, \\ m_t^{-1} = 1 + \frac{S_{21}}{n_1} p_1 \equiv 1 - \frac{p_1}{f_1} = -\frac{\zeta_1}{f_1}, \end{cases}$$

that is,

(185) $$m_t \overset{def}{=} \frac{y_2}{y_1} = -\frac{f_1}{\zeta_1} = -\frac{\zeta_2}{f_2} \quad .$$

These relations can also be directly deduced from the geometry of Fig. 34 (from the similar triangles highlighted with the same type of hachures). Otherwise, by eliminating S_{21}, we also get

(186) $$m_t \overset{def}{=} \frac{y_2}{y_1} = -\frac{n_1}{n_2} \cdot \frac{p_2}{p_1} \quad .$$

Expressions (185), (186) generalize those deduced in the case of spherical diopters, (132), (133).

The focal and principal points and planes completely determine the transformation of paraxial rays by the optical system, a fact based on which they are termed ***cardinal elements***. This is illustrated in Fig. 34, which represents the graphical construction of the image with reference to the four cardinal points F_1, F_2, H_1, H_2 and the two construction rays passing through the foci (***principal rays***). The rule for achieving the construction is simple: any ray coming from the left, which is parallel to the optical axis, is deviated at the intersection with the principal image plane (determined by H_2), and passes through F_2; any ray coming from the left, which passes through F_1, is deviated at the intersection with the principal object planes (determined by H_1), and exits parallel to the optical axis. Obviously, constructing an image using real light rays (which effectively pass through the system) is usually a much more difficult task.



The cardinal points, the object, and the image can also be arranged in a different order than the one considered in Fig. 34. Let us apply the same method of principal rays when the order is that illustrated in Fig. 35, that is: F_1, object, H_2, H_1, F_2. Unlike Fig. 34, where the image is real (the light rays effectively intersect in point Q_2), in this case, the image is virtual, because the rays do not effectively pass through point Q_2, but, upon exiting the system, they behave as though originating in Q_2.

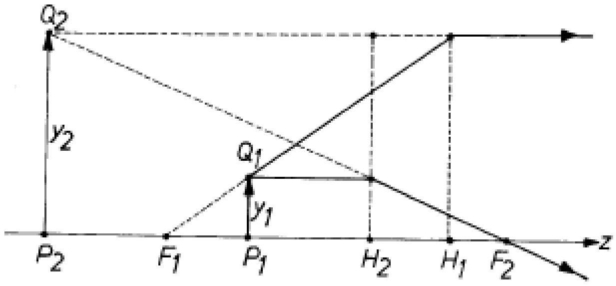


Fig. 35 Another possible situation regarding the positions of the principal planes.

Usually, there is no a priori order of cardinal points, and therefore, if they are all distinct, there are $4! = 24$ different possibilities of arranging them along the optical axis. If we further consider the position of the object, it follows that there are $5! = 120$ different cases of graphical image construction, of which two have been presented in Fig. 34 and 35.

Of course, referring to the properties of the cardinal elements, we could trace the trajectories of any (not necessarily principal) construction rays and construct the image directly, with the help of two arbitrary rays, as illustrated in Fig. 36. Thus, any given incident ray Q_1I_1 may be considered as part of a parallel beam, one of which passes through focus F_1. The focalization point F of the parallel beam with its ray Q_1I_1 is located at the intersection between the ray that passes through F_1 and the image focal plane (determined by F_2). Thus we determine the direction of emergent ray I_2F, conjugate with Q_1I_1. Any other ray J_2G, conjugate with Q_1J_1, and through it, the image Q_2 of Q_1, can be constructed in a similar fashion.

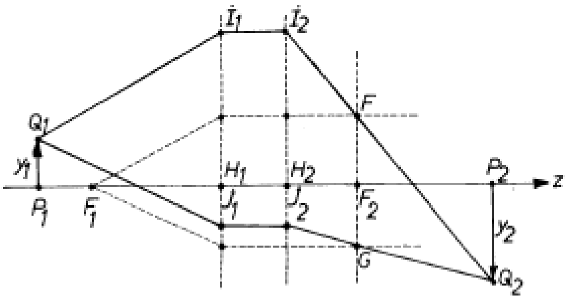


Fig. 36 Constructing the image P_2Q_2 of the object P_1Q_1, in a centered system, when knowing the positions of the focal planes.

Graphical constructions achieved with the help of the cardinal points geometrically demonstrate the relations previously established analytically. Thus, Fig. 37 illustrates how a conical light beam with its apex in point $Q_1(y_1)$, situated within the object focal plane, is transformed into a parallel beam of inclination γ_2, so that $y_1 = -f_2\gamma_2$ (see equation (163)), as well as the opposite phenomenon, the transformation of a parallel beam of inclination γ_1 into a conical beam with its apex in point $Q_2(y_2)$, situated within the image focal plane, so that $y_2 = f_1\gamma_1$ (see equation (164)).

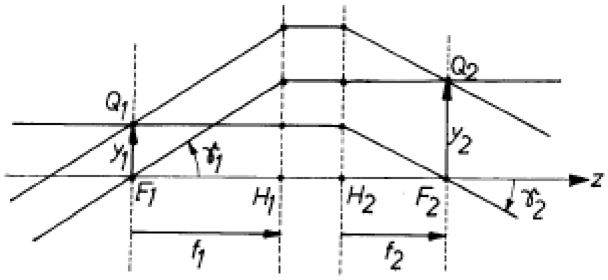


Fig. 37 A conical or parallel beam passing through a centered system (focal planes conjugate with planes located at infinity).

Next, we will give a geometric demonstration of the ***nodal point theorem***, according to which there is a pair of conjugate axial points N_1, N_2, so that for any incidence ray oriented towards the object nodal point N_1, there is a corresponding parallel emerging ray coming from the image nodal point N_2. Let there thus be a conical light beam with its apex at an arbitrary point Q_1 within the object focal plane (see Fig. 38). This beam is transformed by the optical system into one parallel to the ray I_2F_2. From among all the rays originating in point Q_1, only one is parallel to the emerging rays (to I_2F_2), namely the one that propagates along the direction Q_1J_1. This ray intersects the optical axis in N_1, and the corresponding emerging ray J_2N_2 in N_2. However, since triangles $Q_1F_1N_1$ and $I_2H_2F_2$, and $J_1H_1N_1$ and $J_2H_2N_2$ are equal, it follows that $F_1N_1 = f_2$, $N_2F_2 = f_1$, relations independent of the position of point Q_1. In other words, the positions of the nodal points (N_1, N_2) is determined ***only*** by the positions of the foci (F_1, F_2) and the focal distances (f_1, f_2). We can write the relationships thus obtained as

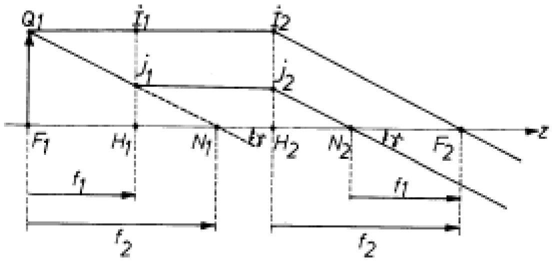


Fig. 38 The nodal points of a centered system (N_1, N_2).

(187)
$$\begin{cases} t_{f_1} - t_{n_1} = f_2, \\ t_{f_2} - t_{n_2} = f_1, \end{cases}$$

where t_{n_1}, t_{n_2} represent the ***positions of the*** (object and image) ***nodal points***. Their expressions can be obtained by introducing the focal distances (equations (165)) and

the positions of the focal points (equations (166)) into equations (187). The nodal points N_1, N_2 can also be used as cardinal points in constructing images.

For convenience, let us review the expressions of the ***cardinal elements***:

$$(188)\quad \begin{cases} f_1 = -n_1/S_{21}, \\ f_2 = -n_2/S_{21}, \\ t_{f_1} = -n_1 S_{22}/S_{21} = f_1 S_{22}, \\ t_{f_2} = -n_2 S_{11}/S_{21} = f_2 S_{11}, \\ t_{p_1} = n_1(1 - S_{22})/S_{21} = f_1(S_{22} - 1), \\ t_{p_2} = n_2(1 - S_{11})/S_{21} = f_2(S_{11} - 1), \\ t_{n_1} = (n_2 - n_1 S_{22})/S_{21} = f_1[S_{22} - n_2/n_1], \\ t_{n_2} = (n_1 - n_2 S_{11})/S_{21} = f_2[S_{11} - n_1/n_2]. \end{cases}$$

All these parameters depend on only three elements (S_{11}, S_{22}, S_{21}) of the transfer matrix S associated with the system (element S_{12} is not independent, since $\det S = S_{11}S_{22} - S_{21}S_{12} = 1$), as well as on the refractive indices of the object (n_1) and image (n_2) space. If $n_1 = n_2$, then $f_1 = f_2$, $t_{p_1} = t_{n_1}$and $t_{p_2} = t_{n_2}$. The focal distances f_1, f_2 only depend on the element S_{21}, which, if the system is immersed in air ($n_1 = n_2 = 1$), effectively represents the convergence of the system, with its sign changed.

Some cardinal points can also be easily determined experimentally. The focal points F_1, F_2 are given by the position of the image of an object located at a distance. Moreover, the nodal points N_1, N_2 can be localized thanks to the property according to which a parallel beam remains focalized in the same point if the optical system is rotated around N_1 and N_2, respectively. In the case of the spherical diopter, equation (145), or of attached diopters (thin lenses), equation (151), we have $S_{11} = S_{22} = 1$, so $t_{f_1} = f_1$, $t_{f_2} = f_2$, and $t_{p_1} = t_{p_2} = 0$, meaning the principal planes coincide with the plane tangent at the diopter vertex and the one tangent at the vertex common to the attached diopters, respectively.

Usually, the equivalent optical plan of any centered optical system, with the help of which, using known rules, we can graphically construct the images, is given by its cardinal points. In turn, these are determined based on the transfer matrix elements, equations (188), calculated based on the data of the actual system. Obviously, a set of cardinal points

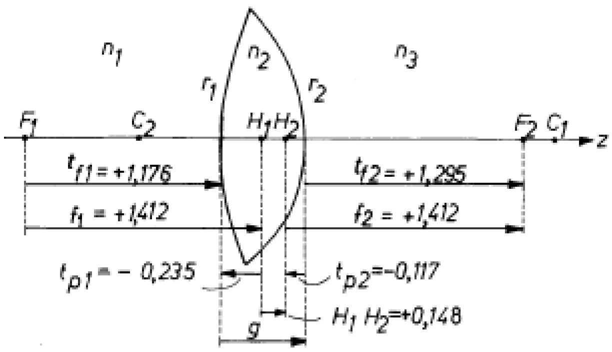


Fig. 39 The cardinal elements of the lens in Fig. 1.29.a.

defines a multitude of different actual systems, which, however, are equivalent regarding the way they construct images.

As a conclusion to this chapter, we will illustrate how the cardinal elements are calculated for three actual optical systems: a simple lens, a compound ocular system composed of two lenses, and a much more complex system, the human eye.

a) Let there be a biconvex glass lens (considered in Fig. 29.a), for which we've already calculated the transfer matrix, equation (152). From equations (188) we readily obtain the cardinal elements illustrated in Fig. 39, in which the numerical data are expressed in centimeters.

b) For the following example, we will consider the ***type Ramsden ocular system***, whose characteristics are given in the following table:

Nr.	n	r (cm)	g (cm)
1	1	∞	0,15
2	1,5	-0,5	0,60
3	1	+0,5	0,15
4	1,5	∞	
5	1		

Calculating matrix $S = R_4 T_3 R_3 T_2 R_2 T_1 R_1$ leads us to the result

$$S = \begin{bmatrix} 0,260 & 0,666 \\ -1,400 & 0,260 \end{bmatrix}$$

According to equations (188), we have

$$t_{f_1} = -S_{22}/S_{21} = +0,186 \ cm,$$

$$t_{f_2} = -S_{11}/S_{21} = +0,186 \ cm,$$

$$t_{p_1} = (1 - S_{22})/S_{21} = -0,528 \ cm,$$

$$t_{p_2} = (1 - S_{11})/S_{21} = -0,528 \ cm,$$

$$f_1 = f_2 = -1/S_{21} = +0,714 \ cm,$$

results represented in Fig. 40. Note that H_1 is located to the right of H_2.

c) The structure and properties of the ***human eye*** impose a series of requirements when designing visual apparatuses. It is a centered, convergent optical



system, containing a series of practically spherical diopters (Fig. 41). The successive transparent media are the ***cornea***, the ***aqueous humor*** ($n = 1.336$), the ***crystalline lens*** and the ***vitreous humor*** ($n = 1.336$). The aperture of the crystalline (the ***pupil***) is controlled by a diaphragm (the ***iris***), and has a diameter of between 2 and 8 mm, varying according to the intensity of light. The optical system of the eye is complicated by the non-homogenous structure of the crystalline. It is a biconvex lens, formed out of around 20,000 successive transparent layers, whose refractive index varies from $n \approx 1.406$, in the case of the peripheral layers, to $n \approx 1.454$ at the center.* By the action of the ***ciliary muscles***, the curvature radii of the ***crystalline*** lens vary within certain limits, the greatest change being effected due to its curved frontal surface.

In this way, the crystalline behaves like a lens of varying focal distance, allowing eye ***accommodation***, so that the image of considered objects may always form on the retina. The eye has the most refraction power at the surface of the cornea (which separates the air from the aqueous humor), but the variations of this power, necessary for accommodation, are caused by the alteration of the crystalline lens' shape. When the ciliary muscles are relaxed, the lens's curvature less pronounced, and its focal distance is at its maximum, so that the image of objects located at greatest distance is formed on the retina (***punctum remotum***). In this case, we say that the eye is not accommodated. In order to focalize on the retina objects located at gradually smaller distances, the crystalline lens must gradually curve, up to a maximum degree of accommodation, corresponding to a ***minimum focusing distance*** (***punctum proximum***), which for a ***normal eye*** is around 25 cm.

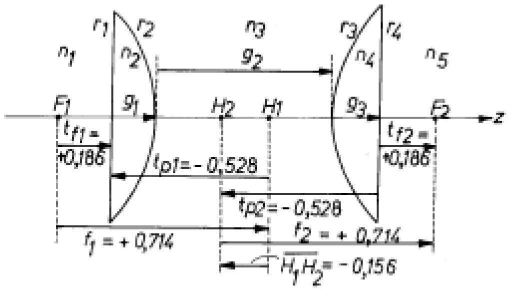


Fig. 40 The cardinal elements of type Ramsden ocular.

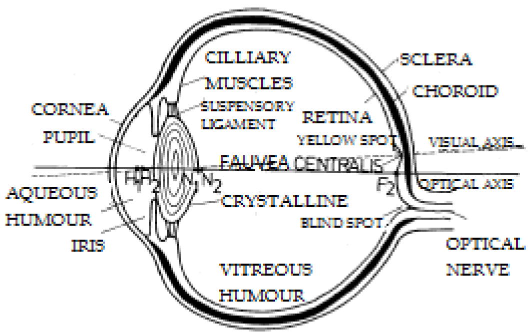


Fig. 41 Structure of the human eye.

Properties similar to a real eye can be achieved if we consider the cornea as a simple refractive surface separating air ($n_1 = 1$) from the aqueous humor (n_2), and the

* Research concerning media of graded (continual) structure with respect to the refractive index is of great actual importance (see Chapter III).

crystalline as a homogenous lens (n_3) immersed between the aqueous humor (n_2) and the vitreous humor (n_4). Such a model for the unaccommodated (relaxed) eye, also called ***Gullstrand's schematic eye***, has the following parameters:

Nr.	n	r (mm)	g (mm)
1	1	+7,8	3,6
2	1,336	-10,0	3,6
3	1,413	-6,0	
4	1,336		

The cardinal elements of this system, according to equation (188), are

$$f_1 = +1,678\ cm; \qquad f_2 = +2,242\ cm;$$

$$t_{f_1} = +1,531\ cm; \qquad t_{f_2} = +1,697\ cm;$$

$$t_{p_1} = -0,147\ cm; \qquad t_{p_2} = -0,545\ cm;$$

$$t_{n_1} = -0,711\ cm; \qquad t_{n2} = +0,019\ cm;$$

Note that $H_1H_2 = N_1N_2 = 0.028\ cm$, but for many applications we may consider that the principal points and the nodal points, respectively, coincide (see Fig. 41). Naturally, the values of the cardinal elements change somewhat compared to the data presented above, when the crystalline becomes accommodated in order to focalize on the retina objects located at a finite distance.

The ***retina***, located at the back of the eye, is the projection screen for the images of external objects. It consists of a mosaic of photosensitive elements, around 130 million "rods" and 7 million "cones" perpendicular to the retina surface and connected to the brain through the optical nerve. The point through which the nerve exits the eye doesn't contain any photoreceptors (the ***blind spot***). The cones are responsible for ***diurnal vision*** and for the sensation of colors, acting like the very fine grain of color film, and the rods are responsible for ***nocturnal vision***. The superficial density of cones increases closer to the ***yellow spot***, or the ***macula***, a depression in the retina, and peaks at around 180.000 cones/mm^2 at its center, across an area of approximately 0.2 mm in diameter, called the ***fovea centralis***. Note that the image of the full moon projected onto the retina has a diameter of approximately 0.2 mm, according to the general formula

(164) $$y_2 = f_1 \gamma_1 \qquad ,$$

(also see Fig. 37), where $f_1 \approx 17\ mm$, as shown in the table presented above. The image of the observed object always forms on the fovea centralis, which furnishes the finest and most detailed information, and determines the ***visual axis*** of the eye, along with the nodal points. Within the foveal area, the medium distance between the centers of neighboring cones is around $2.5 \cdot 10^{-6}\ m$. If we admit the hypothesis that two image points are perceived as distinct if they are separated by at least one unstimulated cone, that is, if we admit that $(y_2)_{min} \approx 5 \cdot 10^{-6}\ m$, it follows that the ***angular resolution limit of the eye*** is

(189) $$(\gamma_1)_{\min} = \frac{(y_2)_{\min}}{f_1} = \frac{5 \cdot 10^{-6}}{1,7 \cdot 10^{-2}} rad = \frac{1}{3400} rad \approx 1' \quad ,$$

a result which well agrees with the physiological estimates regarding ***the normal eye's visual acuity*** ($\stackrel{\text{def}}{=} 1/(\gamma_1)_{min}$).

Other physical and physiological aspects of human eye functioning as an optical receptor

After reviewing the main elements of the structure of the human eye, we would do well to study some of its functional aspects in more depth, as they are of special importance to practical applications.

As is known, because of the geometric aberrations (also see Chapter 2.8) of image forming systems, deviations from stigmatism occur, and the image of an object point is never precisely a point, but a luminous spot which extends spatially to a certain degree. It is therefore natural to ask whether the human eye, as an optical system, presents any astigmatic aberrations, and, if so, to wonder exactly how irksome are they? The answer to the first question is yes, because the image of a point object is ***never*** actually a point. Is it irksome? This time, the answer is no, and the explanation if that in the case of the eye, the image on the retina of the object point is much smaller than $y_{2_{min}} \approx 5\ microns$. That is why the ***visual acuity*** (or the ***separating power***, as it is sometimes called) of the normal eye is determined by the cellular structure* of the retina, and is not influenced by astigmatic aberrations.

* A retinal "cell" is a hexagon inscribed in a circle of diameter $y_2 \simeq 5 \cdot 10^{-6}\ m$; two neighboring "cells" each having in their center a cone are also separated by a cone located between them. According to this model, at least within the foveal area, the retinal structure is similar to that of a honeycomb. Each "cell" sends a single piece of information to the brain (at average cell illumination).

Nor do diffraction phenomena produce any discomfort in the case of visual observation. To better understand this fact, let us calculate the diameter of the diffraction spot (the so called ***Airy*** disk) produced by the pupil, that is, by the eye's diaphragm. If the pupil has the diameter $D = 2R$, the angular opening of the diffraction spot is $\gamma_1 = 1.22\,\lambda/D$. If we consider $\lambda = 550\,nm$ (the midrange of the visible spectrum) and $D = 2\,mm$ (the ***minimum possible*** value of the pupillary diameter), we obtain $\gamma_{1_{max}}^{(diffr.)} = 3.355 \cdot 10^{-4}\,rad \approx 1'09''$, a value expressing the limit of the angular resolution of the eye (see relation (189)). For larger pupillary diameters (up to approximately 8 mm), the angle $\gamma_1^{(diffr.)}$ is smaller than $(\gamma_1)_{min}^{(diffr.)}$, and the diffraction spot forms within a single retina cell. But we've said that details projected onto a single cell are not distinguished, are not separated, for which reason the diffraction spots, generated through diffraction around the pupillary aperture, are not bothersome either.

Since we know the value of $(\gamma_1)_{min}$, it would be interesting to calculate the limit distance $(y_1)_{min}$ between two object points, situated at different distances $|p_1|$ from the eye, that can still be observed as separate. If $|p_1| = 25\,cm$ (the minimum distance of distinct vision, corresponding to the proximal point), we have $(y_1)_{min} = |p_1| \cdot (\gamma_1)_{min} = 0.074\,mm$ and $(y_1)_{min}$ grows as $|p_1|$ grows. In order to distinguish two points located at a distance $y_1 = 1\,mm$ from each other, they must be looked at from a distance smaller than $|p_1|_{max} = 1/(\gamma_1)_{min} = 3.4\,m$.

In order to experimentally determine the eye's visual acuity, multiple procedures and devices can be used: the Landolt ring, the Foucault grating, the Snellen or alphabetical chart, etc. Next we will only refer to the Landolt ring (1874). It is a circular ring, interrupted, like the capital letter C, black on a white background, and its outer (inner) diameter is five (three) times larger than its aperture. The ring, theoretically confirming normal visual acuity, is placed in such a way that the outer angular diameter has a value of $5'$, and the aperture, $1'$.

Visual acuity is easily established based on the distance from which the observer looks at the rig with one eye and can no longer perceive the interruption in the ring.*

The distance $|p_1|$ of the object whose image is projected on the retina following the adjustment of the crystalline curvature radii is termed ***accommodation distance***. For the normal eye, the accommodation distance $|p_1|$ varies between ∞ and $0.25\,m$. For the ***myopic*** eye, when $|p_1| = \infty$, the image forms before the retina (the myopic eye

---

* For details regarding measuring visual acuity, we recommend ***Fiziologie oculară***, Medical Press, Bucharest, 1986 (Chapter XII).

cannot accommodate for infinite distances). For the ***hypermetropic*** eye, for $|p_1| = \infty$, the image forms behind the retina. Myopia can be corrected with eyeglasses with divergent lenses, and hypermetropia, with convergent lenses, which focus the image on the retina.

The limited character of the eye's visual acuity has extremely important consequences and applications, which we should know. Firstly, as we've previously pointed out, the approximately stigmatic character (non-stigmatic, to be precise) of the images given by most instruments, which we have to deal with in practice, does not cause discomfort to the eye (unless the aberrations are great). To better understand the essence of this fact, let us give as an example an instance of viewing images on the television screen. It is known that these images are usually formed out of 625 parallel bands (lines). If the height of the screen is h, the width of a band is evidently $h/625$. If we view the screen from a distance $|p_1|$, each band will be seen at an angle equal to $h/625\ |p_1|$ radians. In order for us not to perceive the discontinuities of the screen, it is necessary for this angle to be smaller than $(\gamma_1)_{min} = 1/3400\ radians$ (in the case of the normal eye). Therefore, it necessarily follows that $|p_1| > 5.44h$. In the case of a regular TV receiver, of $h = 34\ cm$, we get $|p_1| > 1.85\ m$. In reality, for some people, the angular resolution limit $(\gamma_1)_{min}$ can be smaller than the "standard" value of $1'$ (up to $10''$), and for this reason the distance $|p_1|$ must be increased accordingly.

This example shows that there is also an optimum distance from which paintings must be viewed. From a short distance, too many technical details are observable (related to the applying of color to the canvas, etc.), but the overall picture cannot be "seen". On the other hand, when reading or working, a viewing distance shorter than the theoretical distance is preferable, even if this will cause details of no interest to become visible (dust, spots, porous areas on the paper, etc.). As a practical general rule, an angle γ_1 in the interval $(1 \div 5)(\gamma_1)_{min}$ is recommended.

Another practical aspect of great importance is that of ***depth of field***. It is known that, beside the object points situated at the accommodation distance, the eye can also clearly see neighboring points located at smaller or greater distances. The depth of field is the distance between the (upper and lower) limits of this space. Let us consider the eye in Fig. 42 accommodated for distance a and plane (π). Disregarding the geometrical aberrations and diffraction (we've seen why!), let us suppose that the image of point O, in plane (π), is a point P. Then we consider another object point O', situated at a distance a' $(\neq a)$ from the eyes, and consider its image on the retina to correspond to the angle $(\gamma_1)_{min}$. The outer rays of the conical light beam forming this image intersects the accommodation plane (π) forming a circle AB. The spot on the retina is precisely the image of this circle, of diameter $a \cdot (\gamma_1)_{min} \equiv AB$. Based on the

similarity between triangles $O'AB$, $O'MN$, we easily deduce that

$$\frac{D}{a'} = \frac{a(\gamma_1)_{\min}}{\pm(a'-a)} \qquad ((+)\ for\ a' > a;\ (-)\ for\ a' < a) \qquad ,$$

or

$$\frac{1}{a'} = \frac{1}{a} \mp \frac{(\gamma_1)_{\min}}{D} \quad .$$

For the minimum value of the pupillary diameter, $D = 2\ mm$, the last term has the value $(\gamma_1)_{min}/D = 0.147$ diopters, and it immediately follows that

(190) $$a' = \frac{a}{1 \mp 0,147a} \quad .$$

When the eye is accommodated for infinite distances ($a = \infty$), we get $a' = \mp 6.8\ m$ (which is approximated to $7\ m$), where the plus sign corresponds to ***real*** objects, and the minus sign, to ***virtual*** objects. This indicates that a standard, relaxed eye can simultaneously see with sufficient clarity real objects located at distances larger than $7\ m$, and virtual objects located at the same distance, but behind the viewer. (see Fig. 43).

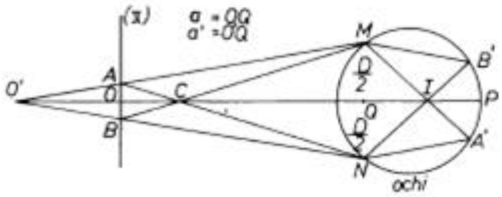


Fig. 42 Determining the depth of field for the human eye.

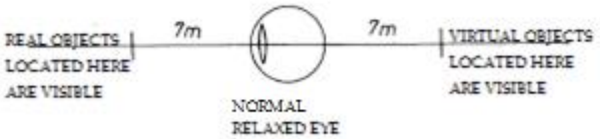


Fig. 43 The space in which the relaxed human eye sees objects (real or virtual).

At maximum accommodation, when $a = 0.25\ m$ (punctum proximum), based on the deduced formula, we get $a' = 24.1\ cm$ and $a' = 26\ cm$, respectively.

When using optical instruments (telescopes, spyglasses, microscope, etc.), the issue of the depth of field must be analyzed separately, because it also depends on the magnification G of the instrument. When using a microscope, for example, the depth of field is practically null (see Chapter 2.5).

Note that the matters discussed here only apply in the situations in which the image is formed in the foveal area of the retina, that is, in the center of the yellow spot, with has a diameter of approximately $0.2\ mm$. Compared to the optical center of the crystalline, the fovea centralis has an angular diameter of about $1°$ (or $0.0175\ radians$). The image of an object situated at a distance $|p_1| = 0.25\ m$, is projected onto the foveal



area if its diameter is 4.35 mm. At a distance $|p_1| = 1\ m$, it fits into the ***optimum visual field***, determined by the angular aperture of the fovea, a disk of diameter 1.74 cm, and so on.

Larger objects, projected onto the rest of the yellow spot, are see less distinctly. Relative to the same optical center, the angular apertures of the yellow spot are 8° horizontally and 6° vertically; this ***less optimal visual field***, at $|p_1| = 0.25\ m$, allows the eye to perceive 3.5 cm horizontally and 2.6 cm vertically; at $|p_1| = 1\ m$, 14 cm and 10.4 cm, respectively, and so on.

Now we can go back to the issue of the optimum distance from which the television screen should be viewed. If we position ourselves so that the image of the screen projected onto the retina fits on the yellow spot, that is, so that the screen height h is seen with an angle of approximately 5°, the width of the bands will be seen at an angle of $5°/625\ lines\ \approx 0.5' < (\gamma_1)_{min}$. This is considered to be the correct position from which one should watch television. The distance $|p_1|$ at which we must position ourselves in order to satisfy the above condition is $|p_1| = 6/0.087 \approx 11.46h$. When $h = 34\ cm$ (a regular screen), we obtain $|p_1| = 3.90\ m$.

If we wish to see all the details of the images on the screen, even with the risk of also seeing the lines, we must move closer, up to about this distance, so that the width of one of the bands is seen at an angle equal to $(\gamma_1)_{min}$. A compromise between the two distances is usually made.

We consider it necessary to also approach the issue of binocular vision, and that of stereoscopic vision*, closely linked to the former, leaving the issue of color perception for later. Note that when we view an object with both eyes, we do not obtain the same information twice. We receive a certain additional amount of information from the overlapping of the two images, processed automatically at the level of the brain.

First, note that the eyes are capable of rotating by about 120° along the vertical plane (the nose-forehead direction), and by 150° within the horizontal plane. This allows for the relatively narrow visual field to be widened due to the mobility of images coming from various visual fields, which are swept by our eyes.

When we view the same object point with both eyes, their visual axes are concurrent in that point, and through their inclination relative to the "forward" direction, the distance to the object is perceived (telemeters work on the same principle). Let us term the distance e between the centers of the eyes ***pupillary distance***; it varies from person to person, and ranges from 54 to 72 mm (we will use

* The capacity to perceive the relative placement of objects in space.

the standard distance of 65 mm). If an object point is located at a distance $|p_1|$, the ***stereoscopic parallax*** η is defined as the ratio $\eta = e/|p_1|$. Let us now consider two object points located at distances $|p_1'|$ and $|p_1''|$. Their parallax difference is evidently $\Delta\eta = e(|p_1''|^{-1} - |p_1'|^{-1})$. It is believed that the parallax difference is perceivable (that is, there is perception of depth of the space in which the two points are seen) if its value is greater than a certain minimum $(\Delta\eta)_{min}$, specific to each person; it varies between $10''$ and $7''$, but it can reach even smaller values ($3''$, for example) through training (as with aviators). For regular analysis, in the case of normal eyes, the standard value $(\Delta\eta)_{min} = 10''$.

Let us once again give several numerical examples. If one of the objects is located at infinity (a star, for example), and we consider the other point at a distance equivalent to $(\Delta\eta)_{min}$, based on the formula for parallax difference, we obtain $|p_1''| = e/(\Delta\eta)_{min} = 65\ mm/10'' = 1340\ m$. Let us use the notation $|p_1|_{min}$. Terrestrial objects located at distances surpassing the value of $|p_1|_{min}$ are perceived as projected on the horizon, its radius having a value of approximately 4 km.

If we write the general relationship expressing the pupillary distance as $e = |p_1|_{min} \cdot (\Delta\eta)_{min}$, we get $(\Delta\eta) = (\Delta\eta)_{min}|p_1|_{min}(|p_1''|^{-1} - |p_1'|^{-1})$, and in the limit case in which $(\Delta\eta) = (\Delta\eta)_{min}$, we have $|p_1|_{min}^{-1} = |p_1''|^{-1} - |p_1'|^{-1}$. Using this relationship, we can calculate the distance $\Delta p_1 = |p_1'| - |p_1''|$ between the point located farther away and the one located closer to the viewer, which can be viewed with the perception of depth. We obtain $\Delta p_1 = |p_1''|^2/(|p_1|_{min} - |p_1''|)$. When $|p_1''| = 25\ cm$ (punctum proximum), it follows that $\Delta p_1 \approx 0.05\ mm$, and when $|p_1''| = \infty$, we have $\Delta p_1 = \infty$. Between the two extreme cases, we obtain intermediary values. For example, for $|p_1''| = 100\ m$, we get $\Delta p_1 \approx 8\ m$.

Let us now analyze how binocular vision can generate the sensation of depth, that is, the stereoscopic effect. In order to do this, we will study the situation outlined in Fig. 44. Let there be two object points M_s and M_d located in the same accommodation plane (π), situated at a distance $|p_1|$ from the eyes. Let us take point M_s to be only observed with the left eye, and point M_d, only with the right. If we extend the visual axes, they intersect in the M points, and so create the impression that there is a single object point M, located at distance $|p|$. Based on the similarity between triangles, we can write $e/|p| = \delta/(|p| - |p_1|)$, meaning $|p| = e \cdot |p_1|/(e - \delta)$. Here $\delta > 0$ when the point M is located farther away than plane (π), the situation illustrated in the left of

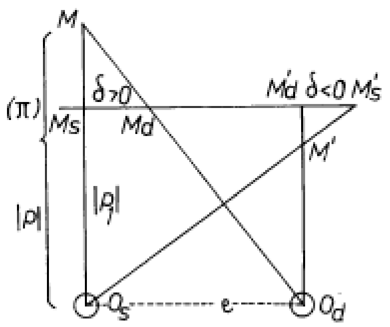


Fig. 44 Binocular vision and obtaining the stereoscopic efect.

the figure, and $\delta < 0$ when point M is located closer than (π), the situation in the right of the figure.

Evidently, in these situations, any plane (π) would not hold only the pair of points M_s and M_d (or M_s' and M_d'), but a multitude of object points paired together, having different values and signs of the parameter δ. The apparent points M and M' corresponding to the various pairs of object points will be seen at different distances $|p|$, before or behind the accommodation plane (π), depending on the corresponding value of the parameter δ. In this case, the sensation of depth is generated at the level of the brain.

If we artificially increase the pupillary distance e (using telemeters or other optical instruments, for example), we can increase the distance $|p_1|_{min} = 1340\ m$, called ***stereoscopic vision threshold***, as well as intensify the stereoscopic effect itself.

The matters heretofore discussed refer only to ***daylight vision***, which features illumination values of at least $100\ lux$, for which the elementary optical receptors are the cones. When illumination drops below $100\ lux$, the cones become less and less sensitive to the luminous stimulus. This is when the rods become important. However, they are not sensitive to color. As we've said before in the first part of this chapter, the number of rods on the retina is much larger than that of cones, but, since they're distributed over a wider area, their superficial density is relatively small.

Let us quickly sketch an analysis of the situation in which illumination is lower than the value $1/100\ \ lux$, that is, an analysis of ***nocturnal*** (***crepuscular***) ***vision***.

There are numerous essential differences between diurnal and nocturnal visions. The first of these is the disappearance of color, determined by the inhibition of cone activity. Secondly, closely connected with this inhibition is decrease in visual acuity and the diminishing of depth perception, while a limit case sees the loss of any stereoscopic effect. In the case of nocturnal vision, the yellow spot and the macula, especially, lose their relative qualities which the rest of the retina lacks. In quite complementary fashion, the peripheral parts of the retina now become more sensitive.

High illumination, necessary for the cones to function, is harmful to the rods, which must be shielded, protected in the case of diurnal vision or very high illumination. This is partly achieved by way of the varying pupillary diameter ($2 - 8\ mm$), a process following which the surface allowing light into the eye decreases in size approximately 16 times. It must be said, however, that this mechanism only accounts for a very small part of the ***adaptation*** interval, whose value is around 10^{12} (the ratio between the maximum and minimum illumination for which the eye can adapt). The basic physiological mechanism of adaptation is of a physicochemical nature, and essentially entails the following. The light incident on a retinal cell effects

the partial decomposition of the ***visual purple*** (or ***rhodopsin***) located there, and the decomposed rhodopsin irritates the visual nerve. Using the energy obtained from the blood supplied to the retina, the rhodopsin is regenerated, and a dynamic equilibrium between decomposed and whole rhodopsin is established. As the light increases in intensity, the equilibrium rests at a smaller quantity of whole rhodopsin, so sensitivity, which is proportional to this quantity, is lessened. Beside this adaptation mechanism, there are the effects produced by the migration of the so-called ***dark pigment***. Exposed to weak illumination, the pigment withdraws towards the liminal space between cells, and spreads across the entire surface of the cell when the illumination to which the eye is exposed intensifies, thus shielding it from its harmful action.

Although it is said that adaptation occurs automatically, we must note that the cumulative processes of all the types of mechanisms does not unfold instantaneously. It has a duration of several (1 – 3) minutes and is asymmetrical: adaptation to intense illumination (entailing the reduction of sensitivity) takes place more quickly than adaptation for weak illumination. Another important aspect is the fact that the fastest mechanism among the three outlined above is that of the variable pupillary diameter.

A very important effect, relevant for the changes that occur upon switching from diurnal to nocturnal vision, was highlighted by Czech physiologist J.E. Purkinje (1825). Let there be two identical surfaces, one red, and the other blue, similarly visible in diurnal conditions, of similar luminance (brightness). As conditions change to those of nocturnal vision and luminosity diminishes, we would notice that the blue surface is much more visible that the red one. The blue surface would appear white, and the red surface, dark gray.

Upon looking into this phenomenon, we have thus approached the issue of color perception. The theoretical foundation regarding colored vision was laid during the last century, by Th. Young, and then by H. Helmholtz. Unfortunately, the associated mechanism is not well understood even today, and there are numerous hypotheses describing its functioning. Still, the essence of the experimentally grounded Young-Helmholtz theory is considered to be valid, for it explains why we clearly perceive colors through diurnal vision, and only differentiate between levels of luminosity through nocturnal vision. According to this theory, the cones are responsible for diurnal vision, and the rods, for nocturnal vision. Moreover, it has been accepted that there are three types of cones on the retina, of different spectral sensitivity, each of which contains a photosensitive pigment for one of the fundamental colors: red, green and blue. When these photosensitive "particles" are stimulated by light, the corresponding color sensation occurs, and through the mixing of sensations in appropriate proportion (depending on spectral sensitivity), all the

other colors are obtained. Based on this observation, it seems perfectly natural to express any color using three "coordinates" signifying the degree to which each "fundamental" color is involved (Maxwell's color triangle), according to the equation "Color = r (red) + g (green-yellow) + b (blue)".

Research conducted during the recent half century have confirmed the suppositions of the Young-Helmholtz theory. For example, studies on spectral absorption undertaken by G. Studnitz in 1941 proved the existence of absorption (and retinal sensitivity) maxima: $\lambda_1 = 470\,nm$ (blue), $\lambda_2 = 550\,nm$ (green-yellow), and $\lambda_3 = 650\,nm$ (red). By calculating the median value of the three absorption curves, we can determine that the sensitivity maximum is $\lambda_d = 555\,nm$. On the other hand, the sensitivity maximum of nocturnal vision, in which rods play the central role, is $\lambda_c = 510\,nm$, meaning it approaches the blue maximum, compared to diurnal vision (see Fig. 45). The Purkinje effect can be easily understood if we take into account these considerations.

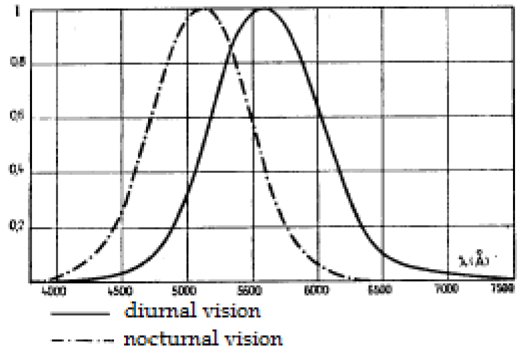


Fig. 45 Sensitivity curves of the human eye.

At the present moment, it is known for certain there is only one type of rod elements, and that they have low chromatic sensitivity (only for $\lambda \leq 625\,nm$). However, we do not know whether each retinal cone has in its structure three specialized chromatic receptors, or if there are three types of cones with distinct spectral sensitivity curves. Based on the experimental data of W.A. Rushton and R.A. Weale (from the 1952 – 1965 period), we may conclude that there are three visual pigments on the retina, each corresponding to a color, with each cone containing a single type of pigment. The research conducted in 1967 by T. Tomita have led to the detection of three groups of cones, with the following distribution: 74% have 611 *nm* sensitivity, 10% have 462 *nm* sensitivity, and 16% have 462 *nm* sensitivity. Therefore, the idea is outlined that there are three types of chromatically specialized cones, each containing a different pigment (***eritrolab***, ***hlorolab***, and tsianolab, respectively)

Although we do not intend to go into details extending beyond the physical scope of this book, we would like to point out that there are people (1% of men and 0.1% of women) whose visual apparatus lack chromatic receptors of a certain kind (***dichromacy***). The most frequently occurring type is ***daltonism***, the absence of cones sensitive to red (the person affected cannot distinguish red from green). It is much rarer for cones corresponding to the other two "fundamental" colors to be absent.

There are also people who lack two of the three types of cones (***monochromacy***), who cannot distinguish colors at all, although such cases are extremely rare. The absence of rods or their presence in insufficient amount is a condition known as ***hemeralopia***. The eye of a person with this condition functions normally in diurnal conditions, but cannot adapt to nocturnal conditions.

2.4 Spherical Lenses

The simplest system of spherical diopters is the spherical lens, formed out of two coaxial diopters with the same internal medium. We will next deduce the general expressions of the transfer matrix S and of the cardinal elements for a general spherical lens (n_2, r_1, r_2, g) immersed in the same medium ($n_1 = n_3$). We have

$$S = R_2 T R_1 = \begin{bmatrix} 1 & 0 \\ \frac{n_2 - n_3}{r_2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \frac{g}{n_2} \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \frac{n_1 - n_2}{r_1} & 1 \end{bmatrix}$$

from which it follows that

(191) $$\begin{cases} S_{11} = 1 + \frac{1-n}{n}\frac{g}{r_1}, \\ S_{22} = 1 + \frac{n-1}{n}\frac{g}{r_2}, \\ S_{12} = \frac{g}{n_1 n}, \\ S_{21} = n_1(1-n)\left(\frac{1}{r_1} - \frac{1}{r_2} + \frac{n-1}{n}\frac{g}{r_1 r_2}\right), \end{cases}$$

where we've used the notation $n = n_2/n_1 = n_2/n_3$ for the relative refractive index of the lens relative to the immersion medium. Note that all the elements of matrix S are linear functions of the lens thickness g.

From equations (188), we obtain the same value for the object and the image focal distance

(192) $$f = f_1 = f_2 = -n_1/S_{21} \quad ,$$

and the same convergence

(193) $$\frac{1}{f} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2} + \frac{n-1}{n}\frac{g}{r_1 r_2}\right) \quad ,$$

an expression which is known as the ***lens maker's formula***. On the enclosure of commercial lenses the value of their convergence in air ($n_1 = n_3 = 1$) is marked.

Depending on the value of the refractive index n, the curvature radii r_1, r_2, and the thickness g, the lens can be positive (convergent lenses), negative (divergent lenses), or null (afocal lenses).

The Huygens-Gauss and the Newton expressions of the conjugate point relationship are written according to equation (182) and equation (183), respectively, that is

(194) $$\frac{1}{p_1}+\frac{1}{p_2}=\frac{1}{f}, \quad \zeta_1\zeta_2=f^2 \quad .$$

Equations (188) and (191) give the positions of the cardinal points

(195) $$\begin{cases} t_{f_1}=fS_{22}=f\left(1+\frac{n-1}{n}\cdot\frac{g}{r_2}\right), \\ t_{f_2}=fS_{11}=f\left(1+\frac{1-n}{n}\cdot\frac{g}{r_1}\right), \\ t_{p_1}=f(S_{22}-1)=\frac{n-1}{n}\cdot\frac{fg}{r_2}, \quad t_{p_1}=t_{n_1}, \\ t_{p_2}=f(S_{11}-1)=\frac{1-n}{n}\cdot\frac{fg}{r_1}, \quad t_{p_2}=t_{n_2}. \end{cases}$$

Usually, the principal points can be located within, as well as outside the lens, and there is no a priori order of their positions. The distance between the principal points of the lens is

(196) $$H_1H_2=g+t_{p_1}+t_{p_2}=g\left[1-f\frac{n-1}{n}\left(\frac{1}{r_1}-\frac{1}{r_2}\right)\right] \quad .$$

An interesting special case is that of the ***afocal lens*** ($S_{21}=0$). The corresponding condition is deduced by eliminating the convergence, equation (193), meaning

(197) $$g=\frac{n}{n-1}(r_1-r_2) \quad .$$

It can be easily verified that this condition is equivalent to that of the coincidence of inner foci of the two diopters of the lens. The angular magnification of the afocal lens (immersed in the same medium) is, according to equations (186) and (191),

$$m_u=S_{22}=1+\frac{n-1}{n}\frac{g}{r_2} \quad ,$$

or, if we introduce the thickness g given by condition (197),

(198) $$m_a = r_1/r_2 \quad .$$

In order to illustrate, let us consider the numerical example $n = 1.5$, $r_1 = 100\ mm$, $r_2 = -2\ mm$, so that from equations (197), (198), we obtain $g = 306\ mm$ and $m_a = -50$, respectively. In this instance, the afocal lens is actually a glass bar with spherical ends, and, constitutes a ***simple telescope***, of significant angular magnification.

A special case of great practical importance is the ***thin lens***. In the case of the ideal thin lens (two completely attached diopters), we have $g \approx 0$, so that the lens maker's formula (193) becomes

(199) $$\frac{1}{f} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad .$$

According to equations (191), the transfer matrix is in this case written as

(200) $$S = \begin{bmatrix} 1 & 0 \\ -n_1/f & 1 \end{bmatrix} \quad ,$$

and the positions of the cardinal points, equations (195), are given by $t_{f_1} = t_{f_2} = f$, $t_{p_1} = t_{p_2} = t_{n_1} = t_{n_2} = 0$. These results are readily generalized for a ***system of attached thin lenses*** (immersed in the same medium), whose transfer matrix is, evidently,

(201) $$S = \prod_{i=1}^{m} S_i = \begin{bmatrix} 1 & 0 \\ -n_1 \sum_{i=1}^{m} \frac{1}{f_i} & 1 \end{bmatrix} \quad .$$

This system is therefore equivalent to a single thin lens whose convergence is equal to the algebraic sum of the convergences of component lenses, that is,

(202) $$\frac{1}{f} = \sum_{i=1}^{m} \frac{1}{f_i} \quad .$$

A better approximation of the formulae for thin lenses, in which thickness g is no longer completely overlooked, can be obtained by introducing the focal distance from equation (199) into equations (195), (196). We thus obtain a fair approximation of the positions of principal points for this kind of real, thin lenses ($fg \ll r_1 r_2$),

(203)
$$\begin{cases} t_{p_1} = \dfrac{g/n}{(r_2/r_1)-1}, & t_{n_1} = t_{p_1}, \\ t_{p_2} = \dfrac{g/n}{(r_1/r_2)-1}, & t_{n_2} = t_{p_2}, \\ H_1H_2 = \dfrac{n-1}{n}g, & N_1N_2 = H_1H_2. \end{cases}$$

The principal points calculated with these formulae in the case of several types of thin glass lenses immersed in air ($n = 1.5$) are illustrated in Fig. 46. In these cases, $H_1H_2 = g/3$ (this relationship also applies to the relatively thick lens presented in Fig. 39).

As an application for a convergent lens ($f > 0$), let us consider the ***simple microscope*** (***the magnifying glass***). In this case, the real object to be investigated ($y_1 > 0$, $p_1 > 0$) is placed near the focus F_1, a little to its right, to be more precise, so that its image is virtual, straight, and magnified ($y_2 > 0$, $p_2 < 0$), as we can see in Fig. 35. In this way, the eye, which is usually located near the image focal plane (F_2), sees the image at an angle θ_2, larger than the angle θ_1, at which it would directly see the object y_1 (that is, without the use of the magnifying glass) if it were placed within the image plane (Fig. 47). According to equation (184), we have

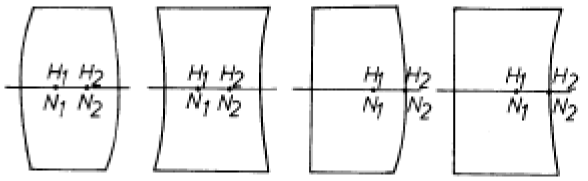


Fig. 46 Principal and nodal points of some types of lens.

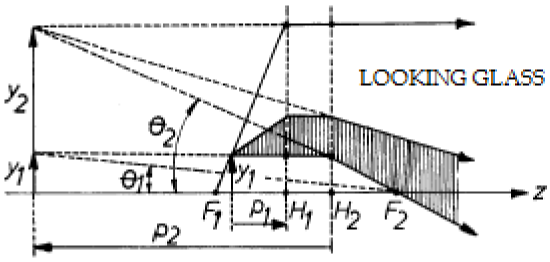


Fig. 47 Image constructed through a looking glass.

(204)
$$m_t \stackrel{def}{=} \frac{y_2}{y_1} = 1 - \frac{p_2}{f} = 1 + \frac{|p_2|}{f} > 0 \quad .$$

The most advantageous distance from which to clearly see the objects or their images is the ***least distance of distinct vision*** (***punctum proximum***), whose values is approximately $0.25\,m = 1/4\ m$. By definition, the ***magnification G*** of the (simple or compound) microscope is

(205) $$G \overset{def}{=} \frac{\theta_2}{\theta_1} = \left(\frac{y_2}{y_1}\right)_{|p_2|=25\,cm} = 1 + \frac{1}{4f}$$, (f is expressed in meters).

Since it is usually the case that $f \ll 25\ cm$, we basically have

(206) $$G \approx 1/4f$$ (f in meters).

The system's magnification can be increased by decreasing its focal distance. However, in practice, it is not possible to go beyond the value $G \approx 8$, corresponding to $f \approx 3\ cm$, if using only one lens, because of the small curvature radii and the corresponding significant geometric and chromatic aberrations. These aberrations can be largely compensated for in appropriately chosen lens systems. An achromatic system of attached lens, with a diaphragm for limiting the light beam, can reach values of $G \approx 25$ ($f \approx 1\ cm$). Aberrations can also be lessened by using a doublet of thin lenses as a magnifying glass, separated by a distance equal to half the sum of their focal distances, such as the Ramsden or the Huygens ocular systems. Usually, they are incorporated into more complex optical systems (microscopes, telescopes), in which they serve to examine the image furnished by the objective. As we will see in the next section, this kind of compound instruments even allow magnification values of $G \approx 10^2 \div 10^3$.

2.5 Compound Systems

Generally, in order to increase magnification and improve the quality of images, two or more lenses are associated. One type of such compound systems are the ***objectives***, which form real images, and the ***oculars***, which form virtual images. The ***microscope*** (intended for observing small, closely grouped objects) and the ***telescope*** (used for observing objects located far away) are, in turn, composed of an objective, which forms the intermediary image, and an ocular, which is used as a magnifying glass, and forms the final virtual image.

A. The Thin Lens Doublet

Let us first consider the simplest of compound systems, namely the ***doublet system, composed of two thin lenses***, of convergences $1/f_1$, $1/f_2$, separated by the distance (the ***doublet thickness***)

(207) $$d = f_1 + f_2 + l \quad ,$$

which are immersed in the same medium of refractive index n (Fig. 48). The parameter l, called ***optical interval***, is the distance between the inner foci of the doublet. Fig. 48 illustrates the construction of the image, by way of the principal rays passing through the foci of the lens components. The first lens, L_1, is called ***field lens***, or ***collecting lens***, and the second, L_2, ***eye lens***. Due to its main application, the thin lens doublet is synonymous with the ***ocular***. The doublets are classified as positive or negative, depending on whether the object focus F_1 of the system is real (that is, is located before L_1) or virtual (that is, located after L_1). Evidently, only the positive ocular can be used as a magnifying glass for observing real objects, and only the negative ocular can be used to observe virtual objects. The convention is for any doublet to be described through its ***symbol***, composed of three algebraic numbers p, q, r, so that

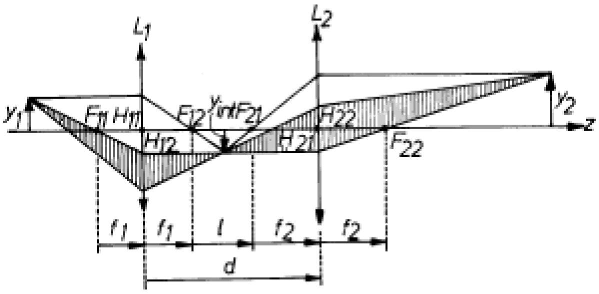


Fig. 48 The thin lens doublet system.

(208) $$\frac{f_1}{p} = \frac{d}{q} = \frac{f_2}{r} = const. \quad ,$$

where, for technical reasons, the constant can be no smaller than a few millimeters. The following table presents several doublets used in practice:

Ramsden	1	1	1	positive
Modified Ramsden	3	2	3	positive
Huygens	4	3	2	negative
Dollond-Huygens	3	2	1	negative
Wollaston	2	3	6	positive

Next we will deduce analytically the formulae of the thin lens doublet. The system's transfer matrix is

$$S = S_2 T S_1 = \begin{bmatrix} 1 & 0 \\ -n/f_2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & d/n \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -n/f_1 & 1 \end{bmatrix} \quad ,$$

and has the elements

(209) $$\begin{cases} S_{11} = 1 - d/f_1, \\ S_{22} = 1 - d/f_2, \\ S_{12} = d/n, \\ S_{21} = -n\left(\frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}\right). \end{cases}$$

Note that all the elements of matrix S are linear functions of the distance d between the lenses, thus allowing the user to suitably adjust the cardinal elements depending on application requirements. From equations (188) and (209), we obtain the expression for the doublet convergence

(210) $$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} = -\frac{l}{f_1 f_2} \quad ,$$

also called the ***Gullstrand formula***, as well as the positions of the cardinal points

(211) $$t_{p_1} = -fd/f_2, \quad t_{p_2} = -fd/f_1 \quad ,$$

(212) $$t_{f_1} = f[1 - d/f_2] = f + t_{p_1}, \quad t_{f_2} = f[1 - d/f_1] = f + tp_2 \quad .$$

Since the immersion media at both ends is the same, we also have $t_{n_1} = t_{p_1}$, $t_{n_2} = t_{p_2}$. We've thus determined all the cardinal elements of the thin lens doublet, namely the focal distance f (equation (210)), the principal points H_1, H_2, through $t_{p_1} = H_1 H_{11}$, $t_{p_2} = H_{22} H_2$ (equation (211)), and the focal points F_1, F_2, through $t_{f_1} = F_1 H_{11}$, $t_{f_2} = H_{22} F_2$ (equation (212)). We can also determine the positions of focal points based on the relations

(213) $$\begin{cases} F_1 F_{11} \equiv t_{f_1} - f_1 = f_1^2/l \ , \\ F_{22} F_2 \equiv t_{f_2} - f_2 = f_2^2/l \ . \end{cases}$$

The latter are nothing other than expressions of the Newton formula, equation (183), the relations between the pair of conjugate points F_1, F_{21} and L_1, and F_{12}, F_2 and L_2.

The simplest oculars, used in microscopes and telescopes, are the ***Huygens*** and the ***Ramsden oculars***, illustrated in Fig. 49. The cardinal elements have been calculated based of equations

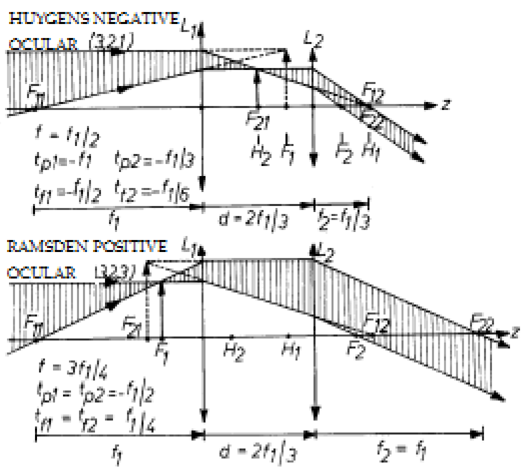


Fig. 49 The Huygens (negative) and the Ramsden (positive) ocular.

(210) – (212), and the images have been constructed by way of the principal rays passing through the foci of the lens components. In both cases, we have considered the eye to be conveniently located behind the ocular, and to view the final (virtual) image (located at infinity) in a relaxed state, which requires the image furnished by an objective system (not shown in the figure) to be located in the ocular system's object focal plane itself, F_1. But, as we have seen, points F_1 and F_{21} are conjugate relative to the field lens L_1 .The image in F_1 represents an object (virtual in the case of the Huygens ocular, real in that of the Ramsden ocular), whose image (real for the Huygens ocular, virtual for the Ramsden version) is formed by the field lens L_1 in the object focal plane F_{21} of the eye lens L_2. Evidently, the final image furnished by lens L_2 (so, by the entire ocular) is virtual, and located at infinity. As in the case of the magnifying glass, this image can be brought closer up to a minimum distance of distinct vision ($\approx 25\ cm$), by moving the ocular appropriately.

As we will see in a later chapter (namely in Chapter 2.7), the condition for convergence achromatization in the case of thin lens doublets made out of the same glass is

(214) $$d = \frac{f_1 + f_2}{2} \quad .$$

It immediately becomes apparent that Huygens-type oculars (4, 3, 2) and (3, 2, 1) satisfy this condition. Moreover, the ocular (3, 2, 1) also has a property that an incidence ray parallel to the principal optical axis is refracted by the same angle on each lens.* This balanced distribution of the refraction effect allows the softening of spherical aberrations. Note that the Ramsden ocular (1, 1, 1) also satisfies equation (214). But in this case, $d = f_2$, so that, unfortunately, focal plane F_{21} and the corresponding image are located right on the field lens L_1. It is for this reason that the viewer also observes through the ocular any impurity or flaw the surface of this lens might have (dust, stains, scratches). This inconvenient aspect is avoided by bringing the two doublet lenses slightly closer, up to a distance of $d = 2f_2/3$, maintaining $f_1 = f_2$. In this way the ***modified Ramsden ocular*** (3, 2, 3) is obtained (presented in Fig. 49). This version no longer strictly satisfies the achromatization condition (214). The relatively weak chromatic aberrations of this ocular are compensated for in the case of the ***Kellner ocular*** (3, 2, 3), in which the only modification consists of replacing the eye lens with a contact achromatic doublet obtained by attaching a convergent lens of

* The condition for this to happen, in paraxial approximation, is $f_1 - f_1 = d$, which, in conjecture with $f_1 + f_2 = 2d$ (equation (214)), results in ocular (3, 2, 1).

crown glass to a divergent lens of flint glass (see Chapter 2.7). Ramsden oculars have the advantage of allowing the introduction of an ***ocular micrometer*** across the F_1focal plane, for measuring the size of the real image formed by the objective system.

No.	n	V*	r (mm)	g (mm)
1	1,0		∞	5,5
2	1,638	55,5	-32,26	1,8
3	1,649	33,8	+36,36	11,5
4	1,638	55,5	-36,36	0,5
5	1,0		+81,30	8,0
6	1,638	55,5	-76,92	0,5
7	1,0		+42,74	12,0
8	1,638	55,5	-42,74	5,4
9	1,720	29,3	+62,89	
10	1,0			

* V is the *Abbe number* (see Chapter 2.7)

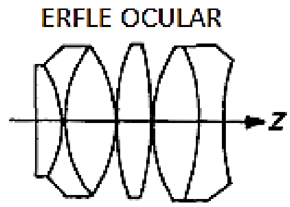


Fig. 50 The Erfle ocular.

No.	n	V	r (mm)	g (mm)
1	1,0		-1,9055	3,39
2	1,620	60,3	-2,8576	0,10
3	1,0		-24,2100	1,60
4	1,620	60,3	-5,4450	0,10
5	1,0		∞	1,80
6	1,611	58,8	-10,1860	0,10
7	1,0		∞	1,40
8	1,620	60,3	-7,0469	0,70
9	1,751	27,8	+7,0469	0,06
10	1,0		+7,4473	1,80
11	1,517	64,5	+100,0000	0,44
12	1,0		-22,9090	3,60
13	1,617	54,9	-13,1830	4,70
14	1,0		-36,9830	3,90
15	1,720	36,2	-18,0300	9,00
16	1,0		+7,5858	3,67
17	1,617	54,9	+5,5463	
18	1,0			

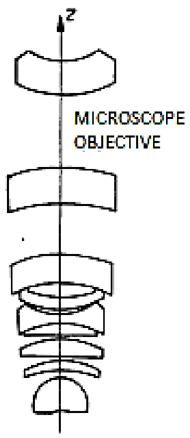


Fig. 51 Microscope objective.

In practice, in order to ensure aplanatism (see Fig. 24) and the correction of tiresome geometric and chromatic aberrations in the desired application, the objectives and oculars of high performance optical systems are themselves fairly complicated compound systems, as in Fig. 50 can be observed the case of an ***Erfle ocular***, with $f = 25.4\,mm$ (Hopkins, 1962), and in Fig. 51, a microscope objective of numerical aperture $n_1 \sin\gamma_1 = 0.85$ and $f = 4.19\,mm$ (Ruben, 1964). It is not our purpose to introduce the reader to the subtle complexities of designing modern dioptric systems, but rather to familiarize him with evaluating, using, and adapting

the systems already widely employed in current practice.

B. Coaxial Optical System Doublet

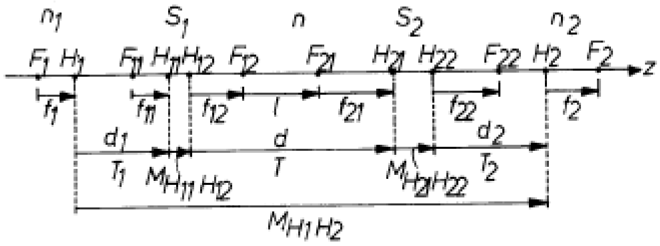


Fig. 52 The cardinal elements of a doublet of coaxial systems.

Heretofore we have learned to calculate the transfer matrix S of any centered optical system, and, with the help of equations (188), to determine its cardinal elements. In this we have obtained the equivalent optical design based on which we may graphically construct the images. So that we may better understand the functioning of microscopes and telescopes, it is necessary that we next generalize the formulae for the thin lens doublet (equations (210) – (213)), to encompass ***any doublet of coaxial optical systems.*** Therefore, let there be system S (see Fig. 52), composed of two subsystems, S_1 and S_2, described by their transfer matrix and, based on equations (188), their cardinal elements, namely

$$S_1 : F_{11}, F_{12}, H_{11}, H_{12}; f_{11}, f_{12},$$

$$S_2 : F_{21}, F_{22}, H_{21}, H_{22}; f_{21}, f_{22}.$$

We will use the notation l for the optical interval, that is, the distance $F_{12}F_{21}$ between the doublet's inner foci, and the notation

(215) $$d = f_{12} + f_{21} + l \quad ,$$

for the doublet thickness, that is, the distance $H_{12}H_{21}$ between the inner principal points. For the sake of generality, we will consider the immersion media to be different. There will therefore be three media: n_1, n_2, and the interior medium n. Our goal is to determine the cardinal elements of this compound system, based on these data. They are

$$S: F_1, F_2, H_1, H_2; f_1, f_2 \quad .$$

The matrix method allows for an elegant analysis of centered optical systems composed of multiple subsystems. Taking advantage of the simple form of the transfer

matrix between principal planes (equation 176), in the case of the present system doublet, we have

$$M_{H_1H_2} = T_2 M_{H_{21}H_{22}} T M_{H_{11}H_{12}} T_1 =$$

$$= \begin{bmatrix} 1 & d_2/n_2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ S_{21}^{(2)} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & d/n \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ S_{21}^{(1)} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & d_1/n_1 \\ 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ S_{21} & 1 \end{bmatrix}$$

After calculating the product of matrices and identification, we obtain

(216) $$S_{21} = S_{21}^{(1)} + S_{21}^{(2)} + S_{21}^{(1)} S_{21}^{(2)} \cdot \frac{d}{n} \quad ,$$

(217) $$\frac{d_1}{n_1} = -\frac{S_{21}^{(2)}}{S_{21}} \cdot \frac{d}{n}, \quad \frac{d_2}{n_2} = -\frac{S_{21}^{(1)}}{S_{21}} \cdot \frac{d}{n} \quad .$$

Next we will replace the 21 elements with the corresponding focal distances, according to their general definition, equations (188), that is

(218) $$\begin{cases} f_1 = -n_1/S_{21}, & f_{11} = -n_1/S_{21}^{(1)}, & f_{21} = n/S_{21}^{(2)} \\ f_2 = -n_2/S_{21}, & f_{12} = -n/S_{21}^{(1)}, & f_{22} = -n_2/S_{21}^{(2)} \end{cases} \quad .$$

From here we obtain the obvious relations

(219) $$\frac{f_1}{f_2} = \frac{n_1}{n_2}, \quad \frac{f_{11}}{f_{12}} = \frac{n_1}{n}, \quad \frac{f_{21}}{f_{22}} = \frac{n}{n_2} \quad .$$

From equation (216), we thus obtain the ***generalization of the Gullstrand formula*** (equation 210), namely

(220) $$\begin{cases} \dfrac{1}{n_2 f_1} \equiv \dfrac{1}{n_1 f_2} = \dfrac{1}{n_2 f_{11}} + \dfrac{1}{n_1 f_{22}} - \dfrac{d}{n f_{11} f_{22}} = -\dfrac{l}{n f_{11} f_{22}}, \\ \dfrac{1}{f_1} = -\dfrac{l}{f_{11} f_{21}}, \qquad \dfrac{1}{f_2} = -\dfrac{l}{f_{12} f_{22}}, \end{cases}$$

equations which allow us to determine the focal distances f_1, f_2. Next, based on equation (217), we obtain the generalized forms of equations (211), (212), namely

(221) $$d_1 = -f_1 d/f_{21}, \quad d_2 = -f_2 d/f_{12} \quad ,$$

relations which allow us to determine the principal points H_1, H_2 from $H_1H_{11} = d_1$, $H_{22}H_2 = d_2$, and the focal points F_1, F_2, from $F_1H_1 = f_1$, $H_2F_2 = f_2$. We can also

determine the positions of the focal points based on

(222) $$\begin{cases} F_1F_{11} = f_1 + d_1 - f_{11} = \dfrac{f_{11}f_{12}}{l}, \\ F_{22}F_2 = f_2 + d_2 - f_{22} = \dfrac{f_{21}f_{22}}{l}, \end{cases}$$

which are generalized versions of equations (213).

We've now analytically determined the cardinal elements of the system doublet. Alternatively, they can also be determined graphically, by the use of the principal rays passing through the foci, as can be seen in Fig. 53. Thus, principal rays of the type $IJF_{12}KLF_2M$, which propagates from left to right, determine the cardinal points F_2, H_2. Similarly, principal rays of the type $MNF_{21}OPF_1I$, which propagates from right to left, determine the cardinal points F_1, H_1. The formulae for the system doublet result directly from the geometry of Fig. 53. Based on the similarity of the triangles with the same hatching, we obtained the relations

(223) $$-f_1/f_{11} = f_{21}/l, \qquad -f_2/f_{22} = f_{12}/l \quad ,$$

(224) $$-f_1/d_1 = f_{21}/d, \qquad -f_2/d_2 = f_{12}/d \quad ,$$

(225) $$F_1F_{11}/f_{11} = f_{12}/l, \qquad F_{22}F_2/f_{22} = f_{21}/l \quad ,$$

which are variant expressions of equations (220), (221), and (222), respectively. Note that relations (222) and (225) are expressions of Newton's formula (equation (183)) for pairs of conjugate points, which in this case are F_1 and F_{21} relative to system S_1, and the pair F_{12}, F_2, relative to system S_2 . Evidently, geometric representations have the advantage of being intuitive.

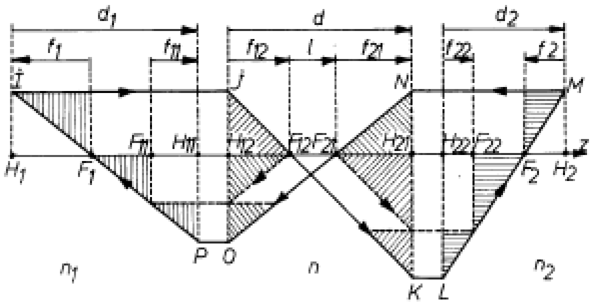


Fig. 53 Image construction in a system doublet and determining the cardinal elements of the global system.

In the special case most often encountered in practice, in which system S_1 and S_2 are immersed in the same medium ($n_1 = n = n_2$), the system of notation is simplified (without the risk of generating confusion, we hope) by effecting the following replacements:

$$f_1 = f_2 \to f,$$

$$f_{11} = f_{12} \to f_1,$$

$$f_{21} = f_{22} \to f_2,$$

so that equations (220), (221), and (222) respectively become

(226) $$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} = -\frac{l}{f_1 f_2} \quad ,$$

(227) $$d_1 = -fd/f_2, \quad d_2 = -fd/f_1 \quad ,$$

(228) $$F_1 F_{11} = f_1^2/l, \quad F_{22} F_2 = f_2^2/l \quad .$$

It is worth notice that equations (226) – (228) of the general doublet of coaxial systems generalize in the same form the equations of the thin lens doublet, that is, equation (210) – (213).

C. Focal and Afocal (Telescopic) Systems

Let us apply the theory elaborated above to the analysis of several systems of practical importance. First, we will consider the class of ***focal systems*** ($S_{21} \neq 0$), designed to form real or virtual images, with high values of transversal linear magnification. We will suppose the media at both ends to be identical, so that equation (184) is written as

(229) $$m_t = \frac{f}{f - p_1} = \frac{f - p_2}{f} \quad .$$

Let us first consider the ***photographic objective***. It is a convergent optical system ($f > 0$) comprising a number of lens. An example would be the Tessar objective (Fig. 29.b), designed to form real images of exterior objects. At most times, the photographic objects are located at a great enough distance so that $p_1 \gg f > 0$, and equation (229) leads to

(230) $$m_t \approx -f/p_1 < 0, \quad |m_t| << 1 \quad .$$



The image is inversed vertically, and the magnification m_t is proportional to the focal distance f of the objective. If the image details are too small, they can no longer be distinguished because of the grained structure of the photographic material. For this reason, objectives with large focal distances must be used. This, however, requires lengthening the photographic camera, since the image basically forms within the image focal plane F_2.

A significant enlarging of the focal distance, for photographing the details of objects located far away, while keeping the photographic camera at a relatively small length, is achieved with the use of the ***telephoto lens*** doublet, composed from the convergent system $S_1(f_1 > 0)$ and the divergent system $S_2(f_2 < 0)$. If we refer to equation (226), the magnification expression (230) becomes

(231) $$m_t \approx f_1 f_2 / l p_1 < 0 \quad ,$$

where f_1, l, $p_1 > 0$, and $f_2 < 0$. This leads to the condition $l = d - f_1 + |f_2| > 0$, meaning $|f_2| > f_1 - d > 0$, which must be met for values $f = f_1|f_2|/l$ as high as possible. Let us illustrate this matter with the numerical example $f_1 = 20\ cm$, $f_2 = -10\ cm$, $d = 15\ cm$, for which, by referring to equations (226) and (227), we obtain $f = 40\ cm$, $d_1 = 60\ cm$, $d_2 = -30\ cm$ (Fig. 54). Note that, although the focal distance of the system is large, the length of the photographic camera, equal to $|f_2|$, remains small (because the cardinal points of the telephoto lens are moved forward a great distance. In practice, the two components S_1, S_2 are achromatic doublets (see Chapter 2.7).

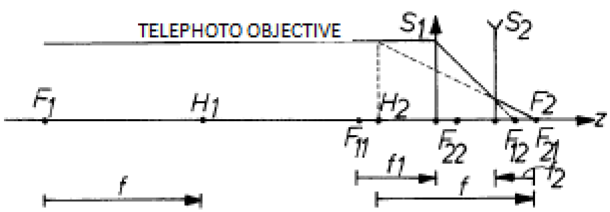


Fig. 54 Simplified plan of a telephoto lens.

Let us next consider the ***compound microscope***, which consists of a convergent objective ($f_1 > 0$) that forms an intermediate inversed image y_{int}, and a convergent ocular ($f_2 > 0$) serving as a magnifying glass, which forms the final virtual image. This principle is demonstrated in Fig. 55, where we've given as example a case in which, in arbitrary units, $f_1 = 17$, $f_2 = 20$, $l = 53$, so $d = f_1 + f_2 + l = 90$; introducing these values into equations (226), (227), or through graphical construction, we get $f \approx -6.5$, $d_1 \approx 29$, $d_2 \approx 34$. Also, for $p_1 = 22$, it follows that $p_2 = -120$ and $m_t \approx -19$. Note that, in order to ensure a numerical aperture $n_1 \sin\gamma_1$ as high as possible, objects must be located very close to the first focal plane (F_{11}) of the objective, basically being

positioned in the first focal plane (F_1) of the system.

Generally, according to equations (226) and (229), we have

(232) $$m_t = 1 - \frac{p_2}{f} = 1 + \frac{|p_2|}{f} = 1 - \frac{l\,|p_2|}{f_1 f_2} \approx -\frac{l\,|p_2|}{f_1 f_2} < 0 \quad ,$$

where the minus sign signifies that the final image is inversed. The approximation made in equation (232) is very useful, since we are only interested in the case in which $p_2/f = l|p_2|/f_1 f_2 \gg 1$. Next, from equations (205), (206), we obtain the ***magnification***

(233) $$G \approx \frac{1}{4f} = -\frac{l}{4 f_1 f_2} \quad , (l, f, f_1, f_2 \text{ in meters}).$$

Compared to the magnification of the simple magnifying glass, formed out of a single lens, ***the magnification of compound microscope*** can be increased to a certain degree, by decreasing the focal distances of the objective and the ocular, and especially by increasing the distance between the inner foci (the standard values for many microscopes are $l = 150 \div 160\ mm$). Thus, for example, a medium magnification $G \approx -160$ is achieved with $f_1 = f_2 = 16\ mm$ and $l = 160\ mm$ ($d = f_1 + f_2 + l = 192\ nn$); in this case, based on equations (226) and (227), we have $f = 1.6\ mm$ and $d_1 = d_2 = 19.2\ mm$. Combinations of various objectives and oculars lead in practice to values within the interval $|G| \approx 25 \div 3000$. Theoretically, the magnification $|G| = 1/4\, f_1 f_2$, equation (233), can reach any value. The ***useful magnification*** is, however, limited by the resolving power of the objective and the observer's eye. Thus, considering that $\theta_2 = G\theta_1 \geq (\gamma_1)^{eye}_{min}$, from equations (123) and (189), we get the ***maximum useful magnification***

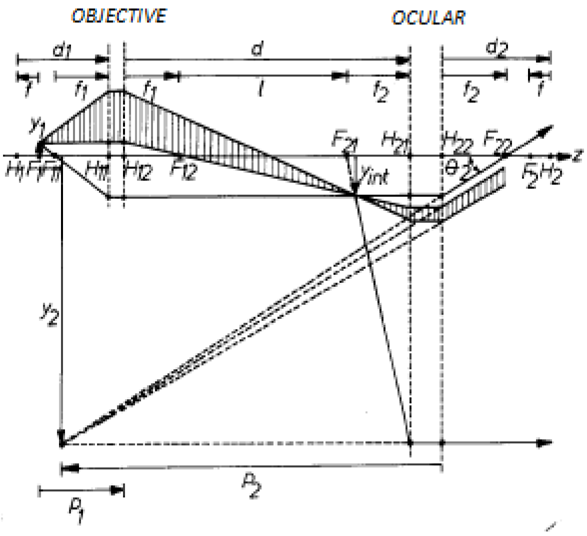


Fig. 55 Image formation in the compound microscope and its cardinal elements.

$$|G| = \frac{|p_2|(\gamma_1)^{eye}_{\min}}{(\delta r_1)_{\min}} = \frac{|p_2|(\gamma_1)^{eye}_{\min}}{0{,}61\lambda_0} n_1 \sin\gamma_1 \quad .$$

By considering $(\gamma_1)^{eye}_{min} \approx 1' \approx (1/3400)_{rad}$, $\lambda_0 \approx 500nm = 5 \cdot 10^{-5}\ cm$, and



$|p_2| = 25\ cm$, we have $|G| \approx 250$ for the dry objective ($n_1 \sin\gamma_1 \approx 1$), and $|G| \approx 375$ for the immersion objective ($n_1 \sin\gamma_1 \approx 1.5$). In order to avoid fatiguing the eye by forcing it to its limit resolving power ($\approx 1'$), a magnification of a value around four times larger is recommended, which leads to the ***useful rule*** $|G| = 1000 \div 1500$. If the magnification excess is too great, there occurs the disadvantage of seeing the diffuse diffraction shapes (the Airy disks) associated to the luminous object points, with no other details.

Unlike the focal systems discussed above ($S_{21} \neq 0$), ***afocal***, or ***telescopic systems*** ($S_{21} = 0$) have infinite focal distances, and their cardinal points are located at infinity (according to equations (188)). As we've seen in Chapter 2.3, these systems have the general property of transforming any parallel beam into another parallel beam (Fig. 32), with an angular magnification

(234) $$m_u = n_1 S_{22}/n_2 = n_1/n_2 S_{11},$$

which has the same value for all conjugate rays, since m_a is a constant of the system. According to the Abbe and Herschel stigmatism conditions (equations (98) and (102)), in paraxial approximation, we also have

(235) $$m_t = n_1/n_2 m_u = S_{11} = 1/S_{22},$$

(236) $$m_a = n_1/n_2 m_u^2 = n_2 S_{11}^2/n_1 = n_2/n_1 S_{22}^2,$$

where we've taken into account the fact that, for afocal systems, $\det S = S_{11}S_{22} = 1$. It follows that the transversal and axial linear magnifications are also constants of the system, meaning they are independent of the object's position. This result is evident in the geometric construction, as is illustrated in Fig. 56 for magnification m_t, with the help of a pair of conjugate rays parallel to the optical axis.

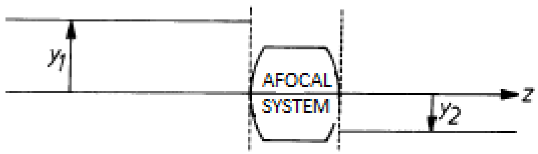


Fig. 56 Simple afocal (telescopic) system.

The expressions for linear magnification (235), (236) ca also be deduced through the matrix method. Thus, if we take into account the defining condition $S_{21} = 0$ for afocal systems, and the expressions (234) for angular magnification, the elements (156) – (159) of the transfer matrix between any two reference planes become

Contemporary
Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017

(237)
$$\begin{cases} M_{11} = S_{11} = n_1/n_2 m_u, \\ M_{22} = S_{22} = n_2 m_u/n_1, \\ M_{21} = S_{21} = 0, \\ M_{12} = \dfrac{S_{11}}{n_1} t_1 + \dfrac{S_{22}}{n_2} t_2 + S_{12} = \dfrac{1}{n_2 m_u} t_1 + \dfrac{m_u}{n_1} t_2 + S_{12}, \end{cases}$$

so that the relationship between conjugate planes $M_{12} = 0$ can also be written as

(238)
$$t_2 = -\frac{n_1 t_1}{n_2 m_u^2} - \frac{n_1 S_{12}}{m_u} \quad .$$

The transfer matrix between conjugate planes of afocal systems thus has two equivalent forms, given by equation (173) and equations (237), respectively, that is,

(239)
$$M_{P_1P_2} = \begin{bmatrix} m_t & 0 \\ 0 & m_t^{-1} \end{bmatrix} = \begin{bmatrix} n_1/n_2 m_u & 0 \\ 0 & n_2 m_u/n_1 \end{bmatrix} \quad ,$$

from which, by identification, results equation (235). Moreover, according to definition $m_a \stackrel{\text{def}}{=} -dt_2/dt_1$, by differentiating equation (238), we obtain equation (236).

Let us next consider the ***afocal doublet*** of coaxial systems (Fig. 57). In this case, the Gullstrand formula (220) for afocal systems ($S_{21} = 0$, that is, $1/f_1 = 1/f_2 = 0$) becomes

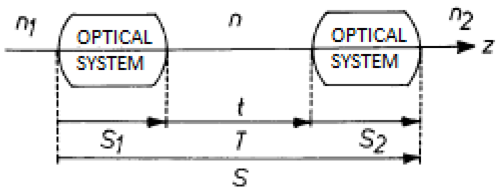


Fig. 57 Afocal doublet of coaxial systems.

$$\frac{n}{n_1} f_{11} + \frac{n}{n_2} f_{22} = d \quad ,$$

or, if we take into account that $f_{11}/f_{12} = n_1/n$, $f_{21}/f_{22} = n/n_2$, equation (219),

$$f_{12} + f_{21} = d \quad ,$$

a condition equivalent to that of coincidence between the inner foci F_{12} and F_{21}.

In order to calculate the magnifications, it is necessary to determine the diagonal elements S_{11}, S_{22}, and to impose the condition $S_{21} = 0$. We will first proceed to calculate the matrix product, according to Fig. 57, whereby

(240) $$S = S_2 T S_1 = \begin{bmatrix} S_{11}^{(2)} & S_{12}^{(2)} \\ S_{21}^{(2)} & S_{22}^{(2)} \end{bmatrix} \cdot \begin{bmatrix} 1 & t/n \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} S_{11}^{(1)} & S_{12}^{(1)} \\ S_{21}^{(1)} & S_{22}^{(1)} \end{bmatrix} ,$$

an equation for which we know that

(241) $$\begin{cases} \det S_1 = S_{11}^{(1)} S_{22}^{(1)} - S_{12}^{(1)} S_{21}^{(1)} = 1, \\ \det S_2 = S_{11}^{(2)} S_{22}^{(2)} - S_{12}^{(2)} S_{21}^{(2)} = 1. \end{cases}$$

We thus obtain the generalized forms of equations (209), as

(242) $$\begin{cases} S_{11} = S_{11}^{(1)} S_{11}^{(2)} + S_{21}^{(1)} S_{12}^{(2)} + (t/n) S_{21}^{(1)} S_{11}^{(2)}, \\ S_{22} = S_{22}^{(1)} S_{22}^{(2)} + S_{12}^{(1)} S_{21}^{(2)} + (t/n) S_{22}^{(1)} S_{21}^{(2)}, \\ S_{12} = S_{12}^{(1)} S_{11}^{(2)} + S_{22}^{(1)} S_{12}^{(2)} + (t/n) S_{22}^{(1)} S_{11}^{(2)}, \\ S_{21} = S_{11}^{(1)} S_{21}^{(2)} + S_{21}^{(1)} S_{22}^{(2)} + (t/n) S_{21}^{(1)} S_{21}^{(2)}, \end{cases}$$

where all the elements of the compound system are linear functions of the reduced distance t/n between the component systems. The condition $S_{21} = 0$ for the compound system to be afocal is satisfied for

(243) $$\frac{t}{n} = -\frac{S_{11}^{(1)} S_{21}^{(2)} + S_{21}^{(1)} S_{22}^{(2)}}{S_{21}^{(1)} S_{21}^{(2)}} .$$

If we introduce this reduced distance in equations (242), take into account equations (241), and replace the 21 elements with the corresponding focal distances, we obtain the matrix for the afocal compound system

(244) $$S = \begin{bmatrix} -f_{21}/f_{12} & \left(f_{12} S_{11}^{(2)} + f_{21} S_{22}^{(1)} \right)/n \\ 0 & -f_{12}/f_{21} \end{bmatrix} .$$

The general expressions for magnifications (234) – (236) in this case become

(245) $$\begin{cases} m_u = \dfrac{n_1}{n_2} S_{22} = -\dfrac{f_{11}}{f_{22}}, \\ m_t = S_{11} = -\dfrac{f_{21}}{f_{12}}, \\ m_a = \dfrac{n_2}{n_1} S_{11}^2 = \dfrac{f_{21} f_{22}}{f_{11} f_{12}}. \end{cases}$$

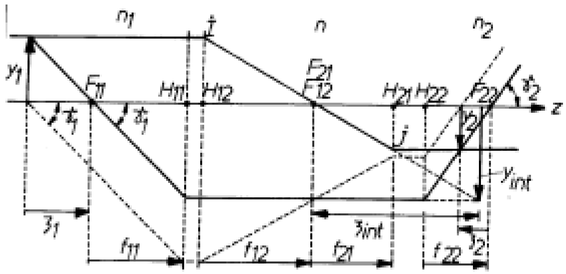


Fig. 58 Image construction in an afocal doublet.

Compared to the analytical method presented above, Fig. 58 illustrates the geometric construction of the image in an afocal doublet with the help of the principal rays passing through the foci F_{11}, $F_{12} = F_{21}$, F_{22} of the two component subsystems S_1 and S_2. First, notice that the similarity between triangles $IH_{12}F_{12}$ and $JH_{21}F_{21}$ means that $m_t \stackrel{\text{def}}{=} y_2/y_1 = -f_{21}/f_{12} = const.$ for any pair of conjugate points. Next, from the Lagrange-Helmholtz invariant $n_1 y_1 \gamma_1 = n_2 y_2 \gamma_2$ (equation (97)), we get $m_u \stackrel{\text{def}}{=} \gamma_2/\gamma_1 = n_1 y_1/n_2 y_2 = -n_1 f_{12}/n_2 f_{21} = -f_{11}/f_{22}$. Finally, if we successively apply Newton's formula, equation (183), we have $\zeta_1 \zeta_{int} = f_{11} f_{12}$, $-\zeta_{int}\zeta_2 = f_{21} f_{22}$, which we differentiate after eliminating ζ_{int} and get $m_a \stackrel{\text{def}}{=} -d\zeta_2/d\zeta_1 = f_{21}f_{22}/f_{11}f_{12}$. Thus, we've used an intuitive geometrical approach to obtain the expression for magnifications (245).

In the special case in which the outer media and the intermediary medium are of the same kind ($n_1 = n = n_2$), we will use the notation $f_{11} = f_{12} = f_1$, $f_{21} = f_{22} = f_2$, so that the formulae (240), (245) for the afocal doublet are simply written as

(246) $$f_1 + f_2 = d,$$

(247) $$m_u = -f_1/f_2, \quad m_t = -f_2/f_1, \quad m_a = (f_2/f_1)^2.$$

Note that these are identical to the formulae for afocal doublet of thin lenses, as we can confirm by using the matrix elements (209).

A relevant application of the afocal doublet is the ***refraction telescope***, an optical instrument designed for observing objects located far away. Like the compound microscope, it consists of a convergent objective S_1 ($f_1 > 0$) which gives and intermediary inversed image y_{int}, and a convergent or divergent ocular S_2, which fulfils the role of a magnifying glass. Due to the very large distance to the observed object, unlike the microscope, the intermediary image is formed on the image focal plane (F_{12}) of an objective of large focal distance. Usually, the telescope functions as an afocal doublet, the mobile ocular being moved until the inner foci coincide ($F_{12} = F_{21}$), so that the eye may view the final virtual image, located at infinity, relaxed (unaccommodated). The following figures illustrate the telescope principle in three versions: the ***astronomical telescope***, or the ***Kepler telescope*** ($f_1 > 0$, $f_2 > 0$, so $m_u <$

0), in Fig. 59.a, the ***Galilei telescope*** ($f_1 > 0$, $f_2 < 0$, so $m_u > 0$), in Fig. 59.b, and the ***terrestrial telescope*** ($f_1 > 0$, $f_v > 0$, $f_2 > 0$, so $m_u > 0$), in Fig. 60, in which, for convenience, the component systems are represented by the thin lenses L_1, L_2, L_v, of focal distances f_1, $|f_2|$, f_v, respectively, whose values give the ratio 7 : 3 : 4. The intermediary, or ***carrier***, lens L_v introduced in the terrestrial telescope is the convergent system that ensures image correction according to equation $m_t = -f_v/\zeta_1 = -\zeta_2/f_v = -1$ (see equation (185)). Note that, at equal angular magnification (in these instances, $|m_u| = 7/3$), the Galilei telescope ($d = f_1 - |f_2|$) is shorter, and the terrestrial telescope ($d = f_1 + f_2 + 4f_v$) is longer than the astronomical telescope ($d = f_1 + f_2$).

Usually, due to the large distances at which the observed objects are located, the useful rays traversing the objective have a very small inclination γ_1 relative to the optical axis. In this case, the most important aberration is the axial chromatic aberration. For this reason, the refraction telescope objectives are basically an achromatic system of attached lenses, usually an achromatic doublet or triplet.

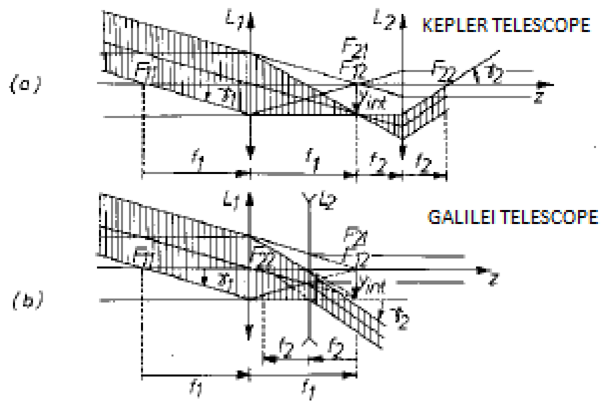


Fig. 59 The Kepler (a) and Galilei (b) telescopes.

Oftentimes, instead of visual observation, photographic recording is preferred. For such applications, a system for forming a final, real image is installed as an extension coaxial with the telescope system, such as the telephoto doublet discussed earlier.

Another interesting application of the afocal doublet is the ***beam expander***, used to enlarge the transverse section of a narrow laser beam. Fig. 59 illustrates how the laser beam is projected along the axis, from right to left. Note that the Galilei doublet (b) is preferable to the Kepler version (a) at high laser powers, in order to avoid ionization and the optical breakdown of the air in the inner, real focus. As can be deduced from the geometry of the figure, the linear rate of expansion of an axial beam (or, in general, of a paraxial one) is equal to the modulus of the angular magnification $|m_u| = |f_1/f_2|$.

D. Triplet Systems

We've heretofore analyzed the properties of forming images with the help of a doublet of coaxial diopter systems. It would be useful to extend the formulae thus obtained to those applicable in the case of triplet systems, and so on, for the additional possibilities of adjusting the distances between component systems that they afford. As in the case of the doublet, the formulae for the ***triplet system*** may be deduced from the corresponding transfer matrix $S = S_3T_2S_2T_1S_1$. Let us illustrate by considering a triplet of thin lenses immersed in the same medium, such as the one presented in Fig. 60. In this case, the Gullstrand convergence formula (equation (210)) extends as

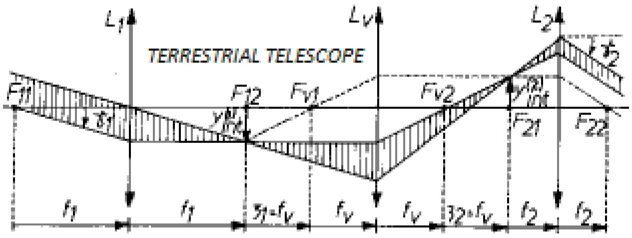


Fig. 60 Terrestrial telescope

(248) $$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3} - \frac{1}{f_1}\left(\frac{1}{f_2} + \frac{1}{f_3}\right)d_{12} - \frac{1}{f_3}\left(\frac{1}{f_1} + \frac{1}{f_2}\right)d_{23} + \frac{d_{12}d_{23}}{f_1f_2f_3} \quad ,$$

where d_{12} and d_{23} are the distances between the first and second lens, and that between the second and third, respectively. We will use this expression to describe the functioning principle of a ***"Zoom" (transfocator) system***. By definition, such a system allows continual variation of its focal distance, so of its magnification, without changing the position of the image. This is most easily achieved in the case of a triplet of thin lenses, by changing the position of the middle lens relative to the other two, which remain fixed. If we use the notation d for the distance between these two lenses, we have $d_{12} = x$ and $d_{23} = d - x$, so that the expression for convergence, equation (248), is in this case written as $1/f = F(x) = Ax^2 + Bx + C$, where A, B, and C are constant. The equation $F(x) = 0$ usually admits two roots, x_1 and x_2, without which the convergence $1/f \neq 0$. In order to achieve large and variable focal distances, the system is designed so that it may allow ranges around these roots. An additional condition is usually imposed, that the roots should be identical, so that moving the middle lens around the position $x_1 = x_2$ should have symmetrical consequences.



E. Reflection (Catoptric) Systems

Unlike the dioptric systems we've heretofore considered, the ***reflectant*** (***catoptric***) ***systems*** have the important advantage of being free of chromatic aberrations. Moreover, for large instruments, it is much easier to produce mirrors than lenses. These advantages were recognized in constructing ***reflection telescopes***, of which the main versions (Newton, Herschel, Gregory, Cassegrain) are illustrated in Fig. 61. A concave parabolic mirror (principal mirror) serves as objective. A small mirror is interposed across the paths of the rays reflected by this objective (secondary mirror), and moves the system's focal plane F into a position convenient for the use of a visual instrument (the ocular), for mounting photographic plates or film, for photoelectric receptors, or for spectral analysis. The only exception is the ***Herschel telescope***, whose focus F is moved directly, through a slight inclination of the principal mirror (of suitably long focal distance). The Newton telescope (also called the side view telescope) employs a plane secondary mirror which deviates the beam at a 90° degree angle to the optical axis. In the case of the ***Gregory telescope***, the focus of the parabolic principal mirror coincides with the first focus of a concave, elliptical secondary mirror, so that the focus F of the entire system forms in its second focus, and the beam exits axially through a small circular orifice cut out in the center of the principal mirror. The ***Cassegrain telescope*** works in a similar fashion, the only difference being that the secondary mirror is hyperbolic and convex. These telescopes use the property of rigorous stigmatism of the Cartesian surfaces of reflection (the paraboloid, ellipsoid, and hyperboloid), as we have seen in Chapter 1.3. The foci of these surfaces are not aplanatic, however, so the angular aperture through which clear images may be obtained only reaches several arc minutes. The ***Schmidt optical system*** (Fig. 61) employs an ingenious solution

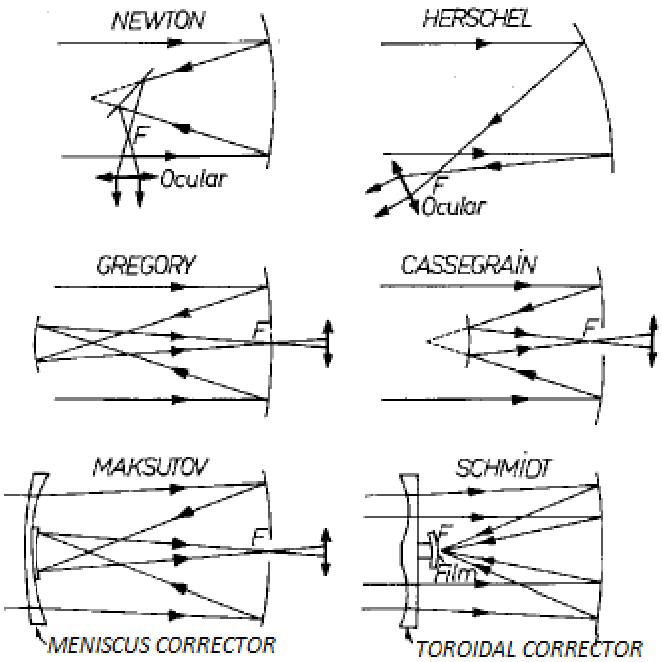


Fig. 61 Main types of reflection telescopes.



to eliminate the spherical aberrations of the spherical (not parabolic) principal mirror, by the use of a suitable correcting refringent blade, whose surface is planar and polished to a slightly toroidal curve. In this way, the marginal and paraxial rays are focused onto the same point F, where the spherical support for photographic film is located. Due to its very large angular aperture, of about 25°, the Schmidt telescope is an ideal instrument for observing the night sky. This system is composed of both reflecting (catoptric) and refringent (dioptric) elements, which places it in the category of ***catadioptric systems***. To this category also belongs the ***Maksutov telescope*** (Fig. 61), which uses a convex or concave secondary mirror, and whose spherical aberrations related to the spherical principal mirror are corrected with the use of a meniscus lens, which is much more easily produced than the aspherical surfaces of the Schmidt correcting blade.

As in the case of the microscope (see the Abbe formula, equation (123)), the diffraction of light imposes a lower limit for the angular aperture γ_1 between the two source points located at great distances (for example, that between the components of a double star) that can still be resolved by the telescope objective, which is

(249) $$(\gamma_1)_{\min} = 1,22\frac{\lambda}{D} \quad ,$$

where D is the diameter of objective lens or mirror. ***The telescope angular resolving power***, defined as $1/(\gamma_1)_{min}$, is thus proportional to D/λ.

Let us illustrate by first considering the eye, which behaves as a refraction telescope objective when viewing objects located far away. If we consider the pupillary diameter of the relaxed eye in daylight to be $D \approx 2\ mm$, and $\lambda \approx 500\ nm$, based on equation (249), we obtain $(\gamma_1)_{min} \approx 1'$. It is remarkable that this value coincides with the physiological limit given by the retinal granular structure, equation (189). This means that the human eye can perceive the maximum of information allowed by the fundamental limit (equation (249)) imposed by the wavelike nature of light. For comparison, the following table gives the diameter of several telescopes and their minimum angular aperture for $\lambda \approx 500\ nm$. The largest diameters are those of the Yerkes telescope, among refraction telescopes, and the Caucasus telescope, from among reflection telescopes.

Objective	D (cm)	$(\gamma_1)_{min}$	Observatory
eye	0.2	$1'$	
lens	12	$1''$	
parabolic mirror	50	$0.24''$	Bucharest



lens	75	0.16″	Pulkovo, Russia
lens	102	0.12″	Yerkes, USA
parabolic mirror	258	0.047″	Mount Wilson, USA
parabolic mirror	508	0.024″	Mount Palomar, USA
parabolic mirror	600	0.020″	Caucasus

In practice, refraction telescopes are used to measure angles and observe the surfaces of planets, while reflection telescopes are used particularly in spectral analysis of light coming from the heavenly bodies. The usable resolving power at low altitudes does not surpass value corresponding to $(\gamma_1)_{min} \approx 0.5''$, because of the refractive index fluctuating along the trajectories of light rays passing through Earth's atmosphere. That is why the location at which the great astronomical observatories are built is chosen based on the optical qualities of its atmosphere. The interference of atmospheric perturbation can only be completely avoided by installing telescopes onboard space shuttles, or on the surface of the Moon.

In theory, the angular magnification $m_u = -f_1/f_2$ (equation (247)) can reach infinitely high values, depending on how great the focal distance f_1 of the objective is, relative to the focal distance f_2 of the ocular. In practice, the length of the telescope is determined by the focal distance f_1 of the objective. The ***useful angular magnification*** of any one telescope is, however, limited by the resolving power of its objective and of the observer's eye. Therefore, considering that $\gamma_2 = m_u \gamma_1 \geq (\gamma_1)_{min}^{eye}$, based on equations (189) and (249), the maximum useful angular magnification is

$$|m_u| = (\gamma_1)_{\min}^{ochi} \ D/1,22\lambda \quad .$$

If we take $(\gamma_1)_{min}^{eye} \approx 1' \approx (1/3{,}400)\ rad$, and $\lambda \approx 5{,}000\ \text{Å} = 5 \cdot 10^{-5}\ cm$, we have $|m_u| \approx 5D\ (cm)$. However, in actual use, the eye suffers considerable strain when observing details at the limit of angular resolution, so that an angular magnification four times greater is recommended, wherefore we obtain the ***useful rule*** $|m_u| \approx 20D\ (cm)$. Further increasing the angular magnification past this value no longer improves the visibility of the object, but only that of the diffuse Airy diffraction disks.

2.6 Diaphragms

In studying centered optical systems, until now we've only approached the issue of image formation, whether they be real or virtual, in convenient positions for recording or visual observation. Other important properties of images, such as



brightness and view field, are dependent on limiting the light beam by the use of ***diaphragms***, that is, of apertures in opaque screens, or of the very borders of the lenses and mirrors that constitute the optical system. For the sake of convenience, we will next analyze the issue of diaphragms with circular aperture, in paraxial approximation.

Let us thus consider a pair of conjugate planes intersecting the optical axis in object point P_1 and image point P_2, respectively. By definition, the ***diaphragm of aperture*** $\boldsymbol{\Delta}$ employed in the system is that diaphragm or those borders of the optical elements (lenses, mirrors) that most limit the beam of light coming from the axial object point P_1 (see Fig. 62 for the general case and Fig. 63 for the special case of a triplet of thin lenses). In order to determine which of the system's diaphragm is the one of aperture corresponding to point P_1 let us first consider the transfer matrix $M_{P_1\Delta}$ between the reference plane containing P_1 and the plane of a certain diaphragm with an aperture of radius R.

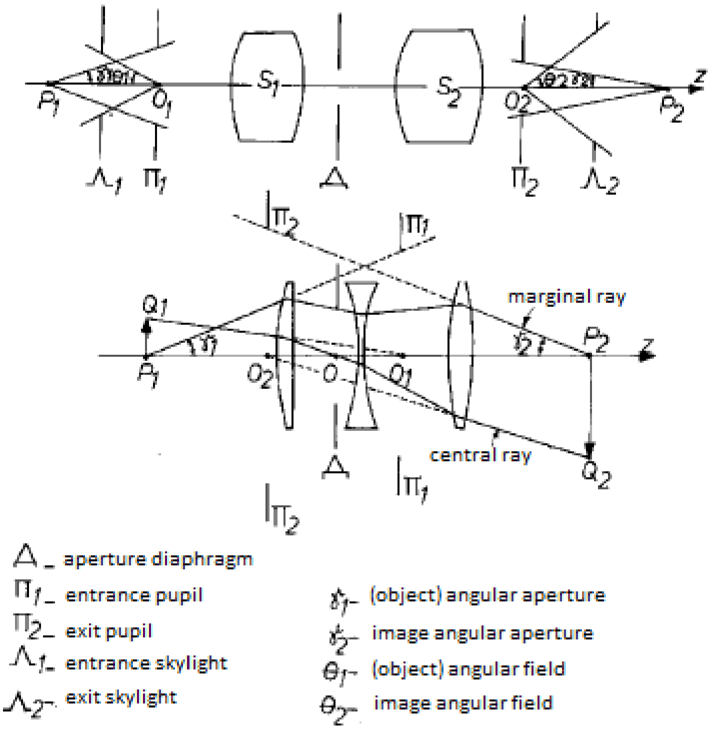


Fig. 62 Aperture diaphragm, pupils, and skylights of a general system.

Fig. 63 Aperture diaphragm, pupils, and skylights of a thin lens triplet.

Transfer equation (161) $y_2 = M_{11}y_1 + M_{12}\Gamma_1$, for a ray passing through P_1 ($y_1 = 0, n_1 = 1, \Gamma_1 = \gamma_1$) and along the aperture's edge ($y_2 = R$) is written as $R = M_{12}\gamma_1$, that is,

$$\gamma_1 = R/M_{12} \quad .$$

According to the definition, it thus follows that the aperture diaphragm is that diaphragm or those lens borders for which γ_1, that is, the ratio R/M_{12}, is smallest. This diaphragm, indicated by the symbol Δ in Fig. 62 and 63, and the corresponding angle γ_1, termed ***angular aperture*** (on the object's side), is particularly important, since it determines the flux of light collected by the system, and therefore image brightness,

as well as its resolving power (see the numerical aperture $n_1 \sin \gamma_1$ of the microscope objective, equation (123)). The image of the aperture diaphragm formed by the front part of the system (system S_1 in Fig. 62 and the first lens in Fig. 63) is called ***entrance pupil*** (Π_1), and the image formed by the hind part (system S_2 in Fig. 62 and the last two lenses in Fig. 63) is termed ***exit pupil*** (Π_2). Obviously, the exit pupil is the image of the entrance pupil formed by the entire system. Note that the angular aperture γ_1 of the cone of light entering the system is determined by the entrance pupil, and the angular aperture (on the side of the image), or the projection angle, γ_2, by the exit pupil.

Another way of determining the entrance pupil, equivalent to the one already described, is to consider the images of all the diaphragms and lens borders formed through the corresponding anterior lenses; the image with the smallest angle $\gamma_1(P_1)$ is the entrance pupil, and the associated physical element is the aperture diaphragm of the system for the point P_1 considered. Alternatively, we can form the images of all the diaphragms and lens borders through the posterior lenses, and determine the exit pupil as the image with the smallest $\gamma_2(P_2)$ angle. Generally, if we change the position of object point P_1 and that of its conjugate image point P_2, another aperture diaphragm, and so another conjugate pair of pupils, could become operative. If the frontal lens or a diaphragm anterior to the frontal lens has a sufficiently small aperture, then it itself is the aperture diaphragm and also the system's entrance pupil. In the case of telescopes, this role is usually played by the borders of the objective lens, so that its image, formed by the ocular, is the system's exit pupil. As we have seen (equation (249)), the diameter D of the objective determines the angular resolving power of the telescope.

Positioning and enlarging the pupils of optical systems has great practical importance. In the case of visual systems, the observer's eye is positioned at the center of the instrument's exit pupil, which must correspond to the entrance pupil of the eye, that is, to the image of the iris's aperture formed by the transparent cornea and the aqueous humor. In order to comfortably align the eye with the instrument, its exit pupil must be slightly larger than the eye's entrance pupil. For example, telescopes designed for daytime observations must have an exit pupil of $3-4\ mm$, and at least $8\ mm$ in the case of nighttime telescopes. In fact, the term "pupil" comes from the requirement that the exit pupil of visual instruments should be approximately equal to the entrance pupil of the eye.

Let us next consider the photographic objective, the telephoto lens, and the telescope objective, designed to form the image of distant objects. In this case, the area of the image is proportional to the square of the objective focal distance, f^2 (according

to $y_2 = -(y_1/p_1)f$, equation (230)). On the other hand, the light flux collected is proportional to the area of the objective's aperture (to the area of the entrance pupil), that is, to D^2. Therefore, the density of the light flux within the image plane varies as $(D/f)^2$. The ratio D/f is termed ***relative aperture***, and its inverse, ***f number***:

$$\text{"}f\ number\text{"} \overset{\text{def}}{=} f/D.$$

Because the photographic exposure time is proportional to $(f/D)^2$, the "f number" is also called ***lens speed***. For example, the "f number" of a lens of focal distance 5 cm and aperture 2.5 cm is thus 2, and is written as $f/2$. The diaphragms of photographic camera objectives are marked with "f numbers", namely 1; 1.4; 2; 2.8; 4; $5.6;\ 8;\ 11;\ 16;\ 22$. Consecutive numbers increase by the multiplying factor $\approx \sqrt{2}$, which indicates a decrease of the relative aperture by the factor $\approx 1/\sqrt{2}$, and a decrease of the light flux density within the image plane by $\approx 1/2$. Thus, for example, the same amount of luminous energy passes through a $f/1.4$ diaphragm in $(1/500)\ s$, through a $f/2$ diaphragm in $(1/250)\ s$, and through a $f/2.8$ diaphragm in $(1/125)\ s$.

The aperture diaphragm and the associated pupils play an important role in the formation of images of spatial objects. Let us illustrate by considering an objective system whose diaphragm aperture of diameter D is the entrance pupil Π_1 (Fig. 64). Let us observe the image P_2 of an axial object point P_1 through a plate of opaque glass as a focalizing screen. If necessary, we will need to place this screen right within the image plane P_2. We move the screen back and forth by a distance of $\pm d\zeta_2$ in order to adjust the image (by looking at it with our bare eyes). However, it is not essential, and will tolerate a ***minimum circle of diffusion*** whose diameter can reach the value $\delta \approx (\gamma_1)_{min}^{eye}\ 250\ mm = (1/3400)\ 250\ mm \approx 0.07\ mm$ without experiencing a significant loss of image clarity.

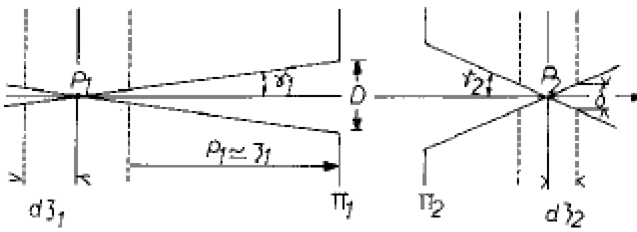


Fig. 64 Defining the circle of minimum diffusion and the (focalization) depth of field.

Evidently, if we use a magnifying glass to adjust the image, the diameter δ of the circle of minimum diffusion will be several times (G) smaller than when doing it with the naked eye. In this case, there will be a tolerance interval, termed ***depth of field*** or ***focalization depth*** in the image space



(250) $$|d\zeta_2| = \frac{\delta}{2\gamma_2} = \frac{f\delta}{D} \quad ,$$

where in order to obtain the last expression we've used the Lagrange-Helmholtz theorem $y_1\gamma_1 = y_2\gamma_2$, equation (97), the expression for magnification $y_2/y_1 = -f/\zeta_1$, equation (183), and we've supposed that the observed object is sufficiently far away so that $\zeta_1 = p_1 - f \approx p_1 (\gg f)$ and $2\gamma_1 p_1 = D$. It thus follows that the depth of field in the image space is inversely proportional to the angular aperture and with the diameter of the entrance pupil. There is a depth of field within the object space $|d\zeta_1|$ corresponding to the depth of field $|d\zeta_2|$ in the image space. If we refer to Newton's formula $\zeta_1\zeta_2 = f^2$ (equation (18)), and to the relation $\zeta_2 \cdot d\zeta_1 + \zeta_1 \cdot d\zeta_2 = 0$, as well as to the expression for $|d\zeta_2|$ above, we get

(251) $$|d\zeta_1| = \frac{|\zeta_1|}{|\zeta_2|}|d\zeta_2| = \frac{\delta}{f} \cdot \frac{p_1^2}{D} \quad .$$

We thus discover that the depth of field in the object space is proportional to the square of the distance to the object, and inversely proportional to the diameter of the entrance pupil (a fact which is qualitatively known by any amateur photographer). In conclusion, the depth of field effect is determined by the finite value of the diameter δ of the circle of minimum diffusion, and by that of the eye's angular resolution $(\gamma_1)_{min}^{eye} \approx 1'$ (equation (189)). This value also indicates the limit to which the aberrations of optical instruments are worth correcting.

The marginal and central rays are especially important to designing optical systems. Rays originating in an axial object point which effectively pass by the borders of the aperture diaphragm are called ***marginal rays*** (Fig. 63). All marginal rays enter the system along a straight line passing by the borders of the entrance pupil and exits the system along a conjugate line passing by the borders of the exit pupil. The ray originating in an extra-axial object point which effectively passes through the center O of the aperture diaphragm is called the ***central ray*** of that point (Fig. 63). This ray enters the optical system along a line passing through the center O_1 of the entrance pupil and exits the system along a conjugate line passing through the center O_2 of the exit pupil. Unlike the aperture diaphragm, which defines the marginal rays and controls the flux of light passing through the system, ***the field diaphragm*** is that diaphragm or those lens borders which most limit the beam of central rays coming from the objective points, and thus controls, as through a window, the system's view field. The image of the field diaphragm formed on the anterior side of the system is termed ***entrance window (skylight)*** (Λ_1), and the one formed on the posterior side,

exit windows (skylight) (Λ_2), Fig. 62. Obviously, just like the pupils, these two windows are located in conjugate positions relative to the entire optical system.

A systematic method of determining the field diaphragm entails forming the images of all the diaphragms and borders through the front lenses; the image seen in the center of the entrance pupil with the smallest angle (the ***acceptance angle***, or the ***object angular field***, θ_1) is the entrance window, and the corresponding physical element is the field diaphragm. Alternatively, we can form the images of all the diaphragms and borders through the posterior lens, and determine the exit window as the image seen in the center of the exit pupil, with the smallest angle (the ***image angular field*** θ_2); the corresponding physical element is the field diaphragm.

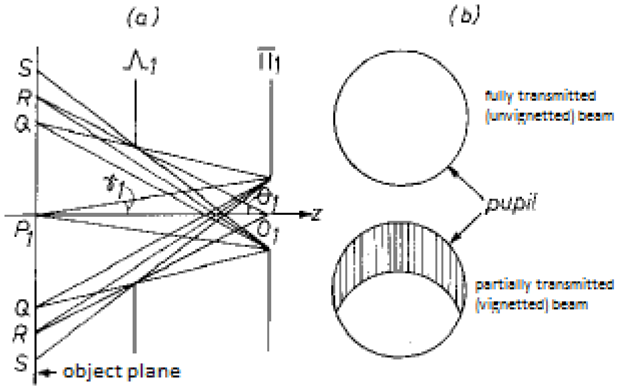


Fig. 65 Defining the view field and highlighting the vignetting effect.

The ***view field*** in any object plane containing the axial point P_1 is defined as the multitude of points within the plan that send light rays into the optical system (Fig. 65.a). These points are therefore situated within the circle of center P_1 and radius P_1S, and are separated into two categories: points analogous to the axial point P_1, such as the extra-axial point Q, which send conical light beams into the system that wholly fill the entrance pupil, and points analogous to the extra-axial point R, whose beams are partially obstructed by the entrance window Λ_1, through the ***window effect***, or ***vignetting*** (Fig. 65.b). The points situated within the circle of radius P_1Q constitute the ***full aperture field***, whose conjugate in the image plane is most luminous, and the points situated within the circular ring between radii P_1Q and P_1S form the ***contour field***, along whose conjugate in the image plane illumination quickly decreases from the its high value on the inner circle to zero on the outer circle. Within the contour field, there is an intermediary circle of radius P_1R (corresponding to the angular field θ_1), along whose conjugate illumination is approximately at half its value in the center of the (total) view field. In conclusion, due to the window effect, the object and image view fields don't have sharply defined borders. To eliminate this inconvenience from optical instruments, the field diaphragm is placed in the plane of an intermediary image, so that the entrance window is located within the object plane, and the exit window, in the image plane. In this way, the entire view field becomes a full aperture field, with almost uniform illumination, and the contour field is reduced to clear

circular borders.

Let us consider in more detail the distribution of the illumination of a real image formed by an optical centered system within the full aperture field. In order to do this, we will suppose that the object is a small plane surface of area dS_1, perpendicular to the optical axis, which radiates according to Lambert's law, that is, its radiance (brightness, illumination) L_1 does not depend on the angle γ (see Chapter 1.3, Fig. 66). The light flux emitted by the axial object point P_1 in the solid angle $d\Omega_{P_1} = 2\pi \sin\gamma \, d\gamma$ is

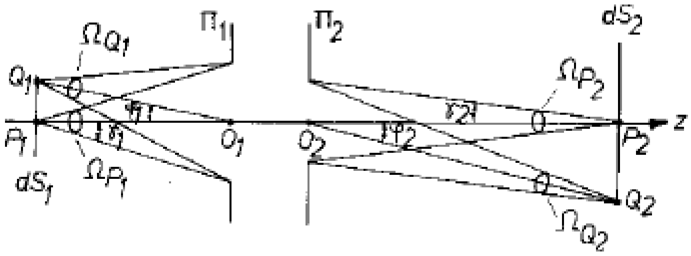


Fig. 66 Determining the illumination of the aplanatic image of a Lambertian source.

(252)
$$d^2F_1 = L_1 d\Omega_{P_1} dS_1 \cos\gamma = 2\pi L_1 dS_1 \cos\gamma \sin\gamma d\gamma \quad .$$

By integrating between $\gamma = 0$ and the angular aperture γ_1, we get the energy flux passing through the entrance pupil, that is,

(253)
$$dF_1 = \pi L_1 dS_1 \sin^2\gamma_1 \quad .$$

We similarly obtain the conjugate energy flux passing through the exit pupil, directed at the axial image point P_2, that is,

(254)
$$dF_2 = \pi L_2 dS_2 \sin^2\gamma_2 \quad .$$

Assuming the optical system satisfies the Abbe sine condition, and that the element of surface dS_2 is the aplanatic image of the element dS_1, we also have the relation (equation (106))

(255)
$$dS_1 n_1^2 \sin^2\gamma_1 = dS_2 n_2^2 \sin^2\gamma_2 \quad .$$

As we've already seen in Chapter 1.3, the energy balance condition $dF_2 = T dF_1$,

where $T \leq 1$ is the transmission (transparency) factor of the system*, along with equations (253) – (255), lead us to the ***Clausius theorem***

(256) $$L_2 = (n_2/n_1)^2 T L_1 \leq (n_2/n_1)^2 L_1 \quad .$$

It follows that, if $n_1 = n_2$, the image radiance cannot surpass the object radiance.

From equations (254), (256), we obtain the ***illumination*** (the density of energy flux) within the image plane

(257) $$E_2 \stackrel{def}{=} dF_2/dS_2 = \pi(n_2/n_1)^2 T L_1 \sin^2\gamma_2 \quad .$$

For sufficiently small angular apertures ($\gamma_2 \ll 1$), the solid angle under which the exit pupil is seen from axial image point P_2 is $\Omega_{P_2} \cong \pi|O_2P_2|^2 \sin^2\gamma_2/|O_2P_2|^2 = \pi \sin^2\gamma_2$. Therefore, equation (257) becomes

(258) $$E_2 = (n_2/n_1)^2 T L_1 \Omega_{P_2} \quad .$$

This expression is valid for illumination in an axial image point P_2. If we repeat the process for an extra-axial image point Q_2, we get

(259) $$E_2 = (n_2/n_1)^2 T L_1 \Omega_{Q_2} \cos\varphi_2 \quad ,$$

where Ω_{Q_2} is the solid angle under which the exit pupil is seen from point Q_2, and φ_2 is the angle between the central ray O_2Q_2 and the optical axis. If we use the symbol Σ for the area of the exit pupil, we have

$$\Omega_{P_2} = \frac{\Sigma}{|O_2P_2|^2} \ , \qquad \Omega_{Q_2} = \frac{\Sigma \cos\varphi_2}{|O_2Q_2|^2} \quad ,$$

from which, if we take into account that $|O_2P_2| = |O_2Q_2| \cos\varphi_2$, we obtain

$$\Omega_{Q_2} = \Omega_{P_2} \cos^3\varphi_2 \quad .$$

---

* We consider T to be independent of λ. The transmission factor in the case of incidence normal or almost normal to air-glass or glass-air separating surfaces is $T = 0.96$. For a thin glass lens immersed in air we thus have $T = 0.96^2 \approx 0.92$. If we also take into account light absorption in optical glass ($A = 0.98$ per centimeter traversed), we conclude that, in the case of a 1 cm thick glass lens immersed in air, we have $T = 0.90$.

Based on this relation, the general expression (259) for the illumination of an aplanatic image of a Lambertian source can now be written as

(260) $$E_2 = (n_2/n_1)^2 T L_1 \Omega_{p_2} \cos^4 \varphi_2 \quad .$$

This "law of fourth power cosine" leads to a quite rapid decrease of image illumination relative to the increase of the angle of field φ_2. In practice, this effect may be countered and image illumination can be homogenized by violating the condition for aplanatism and deliberately introducing coma aberrations. If we increase the angle of field to the value $\varphi_2 \cong \theta_2$, characteristic of the contour field, illumination will decrease drastically due to the vignetting effect.

2.7 Chromatic Aberrations

Until now we've considered the behavior of monochromatic light, so that, in the formulae of geometrical optics, the refractive index ***n*** appears as a single constant pertaining to the material. However, because of the ***light dispersion*** phenomenon, the refractive index depends on the wavelength. For the majority of transparent optical materials, such as inorganic glasses, melted quartz, organic glass (plexiglass), isotropic crystals (fluorine or calcium fluoride, lithium fluoride, sodium chloride, potassium bromide, etc.) or colorless liquids, the refractive index is given in theory by the ***Sellmeier formula***,

(261) $$n^2(\omega) = 1 + \sum_i \frac{A_i}{\Omega_i^2 - \omega^2} \quad ,$$

where $\omega = k_0 c = 2\pi c/\lambda_0$, and A_i, Ω_i are the material constants. In practice, the empirical dispersion formulae are preferred, which depend linearly on the material constants (A, B, C, D), such as the ***Cauchy formula***,

(262) $$n(\lambda_0) = A + \frac{B}{\lambda_0^2} + \frac{C}{\lambda_0^4} \quad ,$$

or, for more exact interpolations, the ***Conrady formula***

(263) $$n(\lambda_0) = A + \frac{B}{\lambda_0} + \frac{C}{\lambda_0^{3,5}} \quad .$$

The ***Herzberger formula*** is an excellent empirical formula used for the wide spectrum between 3,650 Å and 10,000 Å, affording within the visible spectrum a precision of ± 1

to the fifth decimal. It is written as

(264) $$n(\lambda_0) = A + B\lambda_0^2 + \frac{C}{\lambda_0^2 - 2,8 \cdot 10^6} + \frac{D}{(\lambda_0^2 - 2,8 \cdot 10^6)^2} \quad .$$

In general, the refractive indices of optical media within the visible spectrum slowly decrease from blue to red (***normal dispersion***), as is illustrated in the table below for two types of glasses from among the approximately 250 types manufactured by the Schott company. Optical glasses are usually described by two criteria, namely the ***mean refractive index***, n_g, corresponding to a wavelength at the center of the visible spectrum, and the ***mean dispersion*** $\delta_n = n_b - n_r$, corresponding to refractive index variation across a suitable spectrum. For convenience in calculating chromatic aberrations, the dispersion of optical materials is determined using ***Abbe's number***,

(265) $$V \stackrel{\text{def}}{=} \frac{n_y - 1}{n_b - n_r} \approx \frac{n-1}{\delta_n} > 0,$$

where n_b, n_y, n_r are the refractive indices of blue, yellow, and red, corresponding to certain precisely known intense spectral lines, namely:

(266)
the $F(H_\beta)$ hydrogen line, $\lambda_b = 4{,}861.327$ Å,
the $d(D_3)$ helium line, $\lambda_y = 5{,}875.618$ Å,
the $C(H_\alpha)$ hydrogen line, $\lambda_r = 6{,}562.816$.

Spectrum	Light source	Line Fraunhofer name	λ_0(Å)	crown glass BK 7	flint glass SF 11
1 UV	Ar laser		3,511	1.53894	
2 UV	Ar laser		3,638	1.53648	
3 violet	Hg arc	h	4,047	1.53024	1.84211
4 blue	Hg arc	g	4,358	1.52669	1.82518
5 blue	HeCd laser		4,416	1.52611	1.82259
6 blue	Ar laser		4,579	1.52462	1.81596
7 blue	Ar laser		4,658	1.52395	1.81307
8 blue	Ar laser		4,727	1.52339	1.81070
9 blue	Ar laser		4,765	1.52310	1.80945
10 blue	Cd arc	F′	4,800	1.52283	1.80834
11 blue	H arc	F (H_β)	4.861 = λ_b	n_b = 1.52238	1.80645

12 blue	Ar laser		4,880	1.52224	1.80590
13 green	Ar laser		4,965	1.52165	1.80347
14 green	Ar laser		5,017	1.52130	1.80205
15 green	Ar laser		5,145	1.52049	1.79880
16 green	Nd laser		5,320	1.51947	1.79480
17 green	Hg arc	e	5,461	1.51872	1.79190
18 yellow	He arc	d (D_3)	$5{,}876 = \lambda_y$	$n_y = 1.5168$	1.80645
19 yellow	Na arc	D	5,893	1.51673	1.78446
20 red	HeNe arc		6,328	1.51509	1.77862
21 red	Cd arc	C′	6,438	1.51472	1.77734
22 red	H arc	C (H_α)	$6{,}563 = \lambda_r$	$n_r = 1.51432$	1.77599
23 red	ruby laser		6,943	1.51322	1.77231
24 IR	ruby laser		7,860	1.51106	1.76559
25 IR	ruby laser		8,210	1.51037	1.76360
26 IR	GaAlAs laser		8,300	1.51021	1.76312
27 IR	Cs arc	s	8,521	1.50981	1.76202
28 IR	GaAs laser		9,040	1.50894	1.75971
29 IR	Hg arc	t	10,140	1.50731	1.75579
30 IR	Nd laser		10,600	1.50669	1.75444
31 IR	InGaAsP laser		13,000	1.50371	1.74888

(data given in Melles Griot, Optics Guide 3, 1985)

Referring to the table and definition (265), we obtain $V = 64.12$ for BK glass and $V = 25.76$ for SF 11 glass. As V decreases, the δ_n dispersion of the refractive index increases. Note that we may replace the mean refractive index n_y in equation (265) with any value n in the visible spectrum, without the maximum variation of V surpassing around 2%.

Traditionally, optical glasses are separated into two main categories, namely ***flint glass*** (F), for which V varies within the 20 – 50 interval, and ***crown glass*** (K), for which the value varies between 50 and 70. All commercial optical glasses have an index (***catalog code***) made up of two numbers rounded to the third decimal, namely $(n_y - 1)/10V$ (but multiplied by 1,000 for the sake of convenience). Thus, for example, the indices for glasses mentioned in the table above are $785/258$ for super-dense SF 11 flint glass and $517/641$ for BK 7 borosilicate crown glass.

Through the refractive index, the elements of the S matrix and the cardinal elements of optical systems are linked to the wavelength. Obviously, in the case of monochromatic light, such as light filtered through a monochromator or singular frequency laser light, any object gives a single image. If, however, the light is polychromatic, the optical system no longer forms a single image, but a multitude of monochromatic images of different positions and sizes (Fig. 67), so that the image obtained through their overlapping has iridescent borders, and is no longer sharply defined. This inconvenient effect, caused by the dispersion of light, is termed ***chromatic aberration*** or ***chromatism***. We can distinguish between ***axial*** or ***positional chromatic aberrations*** and ***transverse*** or ***size chromatic aberrations***.

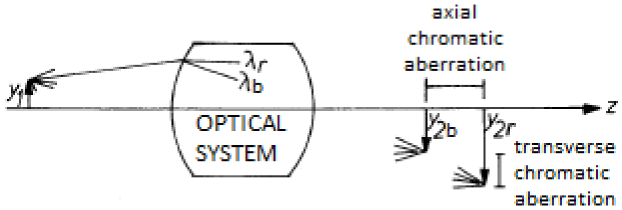


Fig. 67 (Axial and transverse) chromatic aberrations.

These aberrations can be more or less completely eliminated by combining lenses whose contributions to the effect oppose each other. Let us now examine this issue in paraxial approximation. In general, the ***perfect achromatization*** of an optical system ***for two wavelengths*** λ_b and λ_r involve the elimination of all corresponding variations of S matrix elements, that is:

(267)
$$\frac{\delta S_{11}}{\delta n}=\frac{\delta S_{22}}{\delta n}=\frac{\delta S_{21}}{\delta n}=0 \quad ,$$

where $\delta_n = n_b - n_r$. Usually, this system of equations is not compatible, so that (axial and transverse) chromatic aberrations cannot be simultaneously eliminated. That is why, in most practical situations, we have to be satisfied with a ***partial achromatization***, and go in favor of positional or size image precision, depending on the purpose of the instrument.

Let us first consider the simple case of a thin lens. According to equations (199), (200), only element $S_{21} = -n_1/f$ and convergence

(268)
$$\frac{1}{f}=(n-1)\left(\frac{1}{r_1}-\frac{1}{r_2}\right)\equiv K(n-1) \quad ,$$

depend on the refractive index. For brevity, in equation (268) we have used the symbol K to refer to the algebraic sum of the curvatures of diopter components. Evidently, a single thin lens cannot be achromatized because $\delta S_{21}/\delta_n = -Kn_1 \neq 0$. Alternatively,



by differentiating equation (268), we get

(269) $$-\frac{\delta f}{f}=\frac{\delta n}{n-1}=\frac{1}{V}>0 \quad ,$$

or

(270) $$\frac{f_r-f_b}{f_y}=\frac{n_b-n_r}{n_y-1}=\frac{1}{V}>0.$$

Therefore, the axial dispersion $f_r - f_b = f_y/V$ of foci cannot be null, since V is always greater than 0. It is positive for convergent lenses ($f_y > 0$), as is shown in Fig. 68, and negative for divergent lenses ($f_y < 0$). The aberration sign depends on the prismatic shape of convergent or divergent lenses, aberrations becoming thinner or thicker depending on the distance from the optical axis. The extension of the focus axial spectrum is presented in exaggerated proportions in Fig. 68, since, for example, in the case of a type of crown glass of $V = 60$ we have $f_r - f_b = f_y/60$. Also, note that we may replace the mean focal distance f_g in equation (270) with any value f in the visible spectrum.

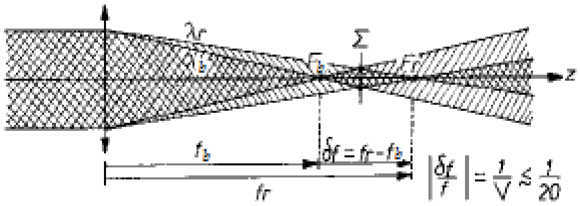


Fig. 68 Axial dispersion $f_r - f_b$ and best "white" image (within plane Σ).

The chromatic aberration is easily discernible in the case of polychromatic light due to the halo surrounding the real image formed on an observation screen. Thus, for example, if we consider a punctual source of "white" light, placed at infinity on the optical axis, a blue dot appears at focus F_b, surrounded by a halo passing into red, and at focus F_r there is a red dot surrounded by a halo passing into blue. The best "white" image appears within an intermediary place (Σ) in the form of a circular disk of minimum diffusion (Fig. 68).

Unlike thin lenses, the focal distance of thick lenses can be achromatized for two wavelengths. In order to do this, there is a condition we may obtain based on equation $\delta(1/f)/\delta n = 0$, equivalent to $\delta S_{21}/\delta n = 0$, and the expression for the convergence of the thick lens, equation (193):

(271) $$g=\frac{n^2}{n^2-1}(r_1-r_2) \quad .$$

Since thickness g is essentially positive, this condition can be met only if $r_1 >$

r_2 . Unfortunately, the other two conditions of system (267), and so, the achromatization of cardinal points, equation (195), can no longer be simultaneously achieved.

Let us study the conditions for achromatization of the thin lens doublet in more detail. We shall start with the condition for achromatization of element S_{21} and for the convergence, as given by the Gullstrand formula

(210) $$\frac{1}{f}=\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2} \quad .$$

From equation $\delta(1/f)/\delta n = 0$, we immediately derive the condition

(272) $$d=\frac{f_1 V_1 + f_2 V_2}{V_1 + V_2} \quad ,$$

in which V_1, V_2 usually correspond to two different types of glass, that is,

(273) $$V_1=\frac{n_{y1}-1}{n_{b1}-n_{r1}}, \qquad V_2=\frac{n_{y2}-1}{n_{b2}-n_{r2}} \quad .$$

In the exceptional case in which the doublet lenses are made out of the same type of glass, that is, when $V_1 = V_2$, equation (272) becomes

(274) $$d=\frac{f_1+f_2}{2} \quad ,$$

a condition which is taken into account when constructing oculars (see Chapter 2.5). As in the case of thick lenses, achromatization is only partial, since dispersion at the cardinal points cannot be simultaneously compensated for. However, the achromatization of focal distance f of the system also involves achromatization of magnification $G = 1/4\,f$ (equation (206)). This apparent achromatization is achieved due to the fact that the images of various colors, although distinct from one another (in terms of position and size), appear to the eye under the same visual angle θ_2, and, thus, in perspective, they ultimately overlap on the retina.

An important method of doublet achromatizing entails attaching together thin lenses made out of different types of glasses, thus obtaining what is known as a ***contact achromatic doublet***. In this case, $d = 0$ and $V_1 \neq V_2$, so that equations (210), (272) become

(275) $$\frac{1}{f_1}+\frac{1}{f_2}=\frac{1}{f} \quad ,$$



(276) $$f_1 V_1 + f_2 V_2 = 0 \quad .$$

Note that in the case of the contact doublet, $S_{11} = S_{22} = 1$, according to equation (201) or equation (209), for $d = 0$, so that the principal planes are not affected by dispersion, and coincide with the plane tangent to the common vertex of component diopters, while the focal planes are fixed by the achromatized focal distance through condition (276). Therefore, the achromatization of the focal distance is the ***perfect achromatization*** with respect to position and size of the contact doublet. If we consider equations (275) and (276), we notice that, in order for the achromatic doublet to not be an instance of the trivial case of null convergence, $1/f = 0$, which would imply that $f_1 = -f_2$, it is necessary that $V_1 \neq V_2$, that is, that the two component lenses be made of different types of glasses. Moreover, since V_1, $V_2 > 0$, it follows that the two focal distances f_1, f_2 must be of different signs, meaning one lens should be convergent, and the other, divergent. It would be interesting to note that, because of the imprecise measuring of refractive indices of his time, inevitable since spectral lines had not yet been discovered, Newton ***erroneously*** concluded that V is the same for all glasses, and so that chromatic aberrations cannot be in principle eliminated (with the exception of the trivial case in which $f_1 = -f_2$). It is for this reason that he devoted his time, successfully, to constructing the reflection telescope, since purely reflective systems would avoid chromatic aberrations (the law of reflection does not refer to the refractive index). The first achromatic doublet was patented by London based optician John Dollond (1758), and made a decisive impact on the perfection of refraction optical systems.

By solving equations (275), (276) for the convergences of the lens components, we obtain the relations necessary for calculating the contact achromatic doublet, namely:

(277) $$\frac{1}{f_1} \equiv (n_{y_1} - 1)\left(\frac{1}{r_{11}} - \frac{1}{r_{12}}\right) = \frac{1}{f} \cdot \frac{V_1}{V_1 - V_2} \quad ,$$

(278) $$\frac{1}{f_2} \equiv (n_{y_2} - 1)\left(\frac{1}{r_{21}} - \frac{1}{r_{22}}\right) = \frac{1}{f} \cdot \frac{V_2}{V_2 - V_1} \quad .$$

Thus, by imposing a certain value for the convergence $1/f$ of the doublet, by choosing the appropriate optical glasses, that is, n_{y_1}, n_{y_2}, V_1, V_2, and referring to the two expressions above, we may calculate the convergences $1/f_1$, $1/f_2$ of the component lenses. In order to avoid values for f_1, f_2 and for curvature radii of the lenses that are too small, the value of the difference $|V_1 - V_2|$ must be sufficiently high. Let us illustrate this procedure by calculating a contact achromatic doublet with $f =$

0.5 m, whose lenses are made from BK 1, 510/635, crown glass, and F2, 620/364, flint glass, respectively. If we introduce the values $1/f = 2\ m^{-1}$ and $V_1 = 63.5$, $V_2 = 36.4$ and refer to the two expressions above, we get $1/f_1 = 4.686\ m^{-1}$ and $1/f_2 = -2.686\ m^{-1}$. Obviously, the sum of component lens convergences must be equal to the convergence of the doublet, equation (275).

Usually, the adjacent surfaces of the contact doublet may not have the same ray curvature, in which case contact is achieved either at the center only, or only at the borders. For any given f_1, f_2, n_{y_1}, n_{y_2}, the first expressions in (277) and (278) are two relations between four ray curvatures, so two of them can be randomly chosen. In practice, this considerable liberty is used to minimize geometrical spherical and coma aberrations, taking advantage of the fact that their effects in the case of convergent and divergent lenses are of opposite sign.

A simple achromatic doublet is the ***Fraunhofer achromat*** (Fig. 69), composed of an equi-convex crown lens ($r_{12} = -r_{11}$), in complete contact ($r_{12} = r_{21}$) with a flint, basically planar-concave lens. The two elements of this doublet are usually bound with a transparent adhesive (such as polyester). Using the numerical data from the above example ($1/f_1 = 4.686\ m^{-1}$, $1/f_2 = -2.686\ m^{-1}$, $n_{y_1} = 1.510$, $n_{y_2} = 1.620$) and the imposed relations between ray curvatures, referring to the first expressions in (277), (278), we get $r_{11} = -r_{12} = -r_{21} = 21.8\ cm$ and $r_{22} = -381.9\ cm$. It is usually recommended that the frontal lens should be the one made out of crown glass, because of its higher resistance to wear.

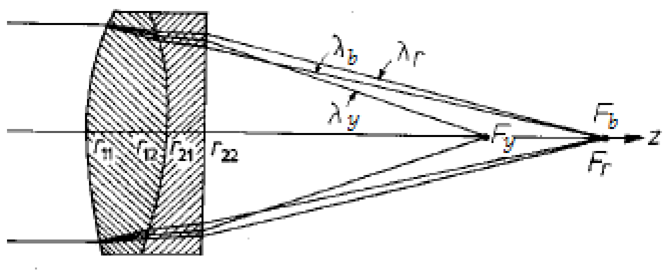


Fig. 69 The Fraunhofer achromat.

We have heretofore considered achromatization conditions for only two wavelengths λ_b, λ_r. In the case of the contact doublet, this implies equality between the corresponding focal distances, $f_b = f_r$ and coincidence of foci F_b, F_r (Fig. 69). However, for other wavelengths λ_x, the corresponding focal distance f_x deviates from the value $f_b = f_r$, imposed for achromatization, and so represents the so-called ***secondary spectrum***, or ***residual chromatic aberration***. In order to determine the distribution of foci F_x for this spectrum, let us differentiate equation (275), and use equation (269) on the component lenses, that is,

(279) $$-\frac{\delta f}{f^2} = -\frac{\delta f_1}{f_1^2} - \frac{\delta f_2}{f_2^2} = \frac{1}{f_1} \cdot \frac{\delta n_1}{n_1 - 1} + \frac{1}{f_2} \cdot \frac{\delta n_2}{n_2 - 1},$$

in which case we will this time consider $\delta f = f_b - f_x$ and $\delta n = n_b - n_x$. Therefore,

(280) $$\frac{f_x - f_b}{f^2} = \frac{1}{f_1} \cdot \frac{n_{b_1} - n_{x_1}}{n_{b_1} - 1} + \frac{1}{f_2} \cdot \frac{n_{b_2} - n_{x_2}}{n_{b_2} - 1} = \frac{1}{f_1 V_{x_1}} + \frac{1}{f_2 V_{x_2}} = \frac{P_{x_1}}{f_1 V_1} + \frac{P_{x_2}}{f_2 V_2} \quad ,$$

in which, for the sake of convenience, we've introduced the ***modified Abbe number***:

(281) $$V_x \overset{def}{=} \frac{n_b - 1}{n_b - n_x} \quad ,$$

and the ***relative partial dispersion:***

(282) $$P_x \overset{def}{=} \frac{V}{V_x} = \frac{n_b - n_x}{n_b - n_r} \quad .$$

If we refer to equations (277) and (278), we may rewrite the last expression in equation (280) as

(283) $$\frac{f_x - f_b}{f} = \frac{P_{x_1} - P_{x_2}}{V_1 - V_2} \quad ,$$

in which form it constitutes the ***secondary spectrum equation***. This equation allows calculating the relative difference between focal distance f_x and the achromatization focal distance $f_b = f_r$, with the help of partial dispersions P_{x_1}, P_{x_2}, and depending on wavelength λ_x. Figure 70 illustrates these calculations for the contact achromatic doublet made out of BK 7 crown (lens 1) and SF 11 flint glass (lens 2), using the refractive indices listed in the table at the beginning of this chapter. The secondary spectrum appears folded onto itself, with a minimum focal distance f_{min} corresponding to the radiation of wavelength λ_m, in the vicinity of yellow line D_3 (λ_y). With the exception of f_{min}, foci corresponding to the various monochromatic radiations coincide two by two. What is remarkable is that the deviation of foci in the visible spectrum from the achromatization value $f_b = f_r$ may be neglected in

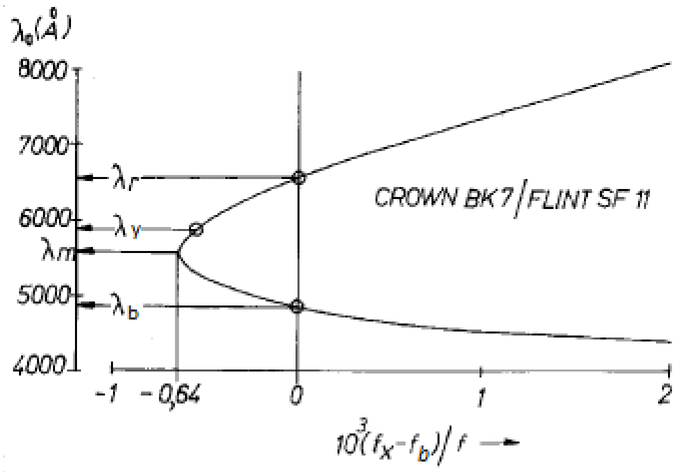


Fig. 70 An analytical representation of the secondary spectrum equation for the BK 7 crown/SF 11 flint doublet.

numerous applications. Indeed, based on Fig. 70, we may conclude, for example, that $|f_y - f_b|/f \approx 1/1800$ (the minimum deviation within the visible spectrum of commercial achromatic doublets is around $1/2000$). Considering this result, it is obvious that the extension of the secondary spectrum suggested by Fig. 69 is greatly exaggerated. For comparison with this performance of the achromatic doublet, let us recall that the relative extension of the spectrum of the foci of a single lens is $(f_r - f_a)/f_y = 1/V$, equation (270), whose value is greater by a multiplier of the order of tens.

According to equation (283), reduction of the secondary spectrum can be achieved, in principle, by choosing a pair of glasses whose difference of partial dispersions $P_{x_1} - P_{x_2}$ is as small as possible, and/or whose difference of Abbe numbers $V_1 - V_2$ is as great as possible. Unfortunately, the first method proves impracticable, since existing optical glasses seem to have a P difference approximately proportional to their V difference, so that selecting glasses does not significantly influence the extension of the secondary spectrum. The second method is more useful. It involves associating a glass of small V with the fluorine crystal, which has a very large value for V ($V = 95.4$). This last method, as we have seen (equations (277) and (278)), also has the additional advantage of the focal distances f_1, f_2, and so of the lenses' rays curvatures, being relatively large, so that the refraction angles and geometrical aberrations are relatively small. By using both methods described above, the glass and fluorine doublet can by achromatized ***for three wavelengths***, λ_b, λ_r, λ_x, ($f_b = f_r = f_x$).

The ***thin lens triplet***, made out of three different glass types, is much more easily achromatized ***for three wavelengths***, λ_b, λ_r, λ_x, or ***for four wavelengths***, λ_b, λ_r, λ_x, λ_z. In the last instance, the system is called ***superachromat***, and it practically allows the complete annihilation of the secondary spectrum in the entire visible domain, and in the near infrared and ultraviolet domains. Let us lastly determine the conditions for the achromatization of the ***superachromatic triplet***. We start from the convergence expression:

(284)
$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3} \quad ,$$

and then impose the equalities $f_b = f_r = f_x = f_y$. By differentiating equation (284), we get:

(285)
$$\frac{\delta f_1}{f_1^2} + \frac{\delta f_2}{f_2^2} + \frac{\delta f_3}{f_3^2} = 0 \quad ,$$

or, if we use equation (269) for component lenses:

(286) $$\frac{1}{f_1}\cdot\frac{\delta n_1}{n_1-1}+\frac{1}{f_2}\cdot\frac{\delta n_2}{n_2-1}+\frac{1}{f_3}\cdot\frac{\delta n_3}{n_3-1}=0 \quad .$$

For $\delta n = n_b - n_r$, we get:

(287) $$\frac{1}{f_1V_1}+\frac{1}{f_2V_2}+\frac{1}{f_3V_3}=0 \quad ,$$

an equation which for the triplet generalizes condition (276). Similarly, for $\delta n = n_b - n_x$, we have:

(288) $$\begin{cases} \dfrac{1}{f_1V_{x_1}}+\dfrac{1}{f_2V_{x_2}}+\dfrac{1}{f_3V_{x_3}}=0, \\ \dfrac{P_{x_1}}{f_1V_1}+\dfrac{P_{x_2}}{f_2V_2}+\dfrac{P_{x_3}}{f_3V_3}=0, \end{cases}$$

and for $\delta n = n_b - n_z$,

(289) $$\begin{cases} \dfrac{1}{f_1V_{z_1}}+\dfrac{1}{f_2V_{z_2}}+\dfrac{1}{f_3V_{z_3}}=0, \\ \dfrac{P_{z_1}}{f_1V_1}+\dfrac{P_{z_2}}{f_2V_2}+\dfrac{P_{z_3}}{f_3V_3}=0. \end{cases}$$

Conditions (287), (288), (289) can be simultaneously satisfied by any triplet of glasses that, in graph (P_x, P_z), are positioned along a straight line, that is, for which

(290) $$P_{z_1}=aP_{x_1}+b, \quad P_{z_2}=aP_{x_2}+b, \quad P_{z_3}=aP_{x_3}+b \quad ,$$

where a, b are constants. Many triplets composed of the glasses available today satisfy this condition.

2.8 Geometrical Aberrations

We've considered centered optical systems in paraxial approximation, so that for any object point Q_1 there is a corresponding conjugate image point Q_2. In other words, within the paraxial domain, the optical system transforms a conical light beam with its apex in Q_1, into a conical light beam with its apex in Q_2, and, respectively, a spherical wave with its center in Q_1 into a spherical wave with its center in Q_2. If,

however, the light beam originating in object point Q_1 is no longer paraxial, the corresponding emerging rays no longer converge in the paraxial (Gaussian) punctual image Q_2, but pierce the Gaussian image plane in various points Q_2^*, and, respectively, the emerging wave deviates from its ideal spherical shape (Fig. 71). These deviations from the ideal punctual image, caused by extra-paraxial rays, are termed ***geometrical aberrations***, or ***monochromatic aberrations***, since they appear even when the light is perfectly monochromatic.

Because of geometrical aberrations, every object point Q_1 corresponds to a diffuse light spot within the Gaussian image plane Q_2, which is limited by an ***aberration curve***, determined by marginal rays. Usually, geometric deviations from the paraxial image Q_2 can be described quantitatively using ***ray aberration vectors***, $\overrightarrow{Q_2Q_2^*}$, and the path difference δ, termed ***wave aberration***, between a ***real (deformed) wavefront*** and a ***reference Gaussian sphere***, with its center in the paraxial image point Q_2. To better understand these notions, let us consider the real wavefront and the Gaussian reference sphere passing through the O_2 center of exit pupil Π_2 (Fig. 71). It will next be advantageous to consider a Cartesian system of coordinates Q_2xyz, with its origin in Q_2, and with the Q_2z axis oriented along direction O_2Q_2, so that the equation for the reference sphere is simply written as

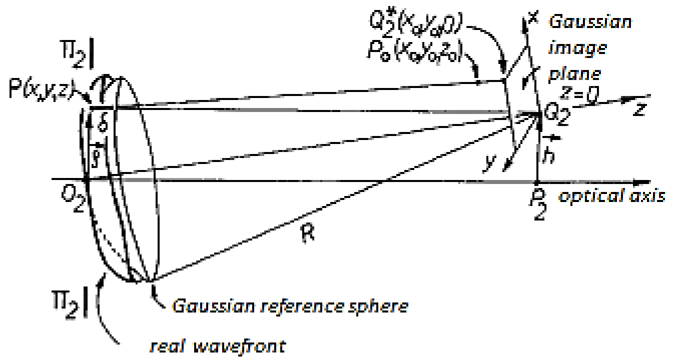


Fig. 71 *Gaussian reference sphere and the actual wavefront*

(291) $$x^2+y^2+z^2=R^2 \quad ,$$

where $R=|O_2Q_2|$. Next, we will consider the real wavefront to deviate from this sphere, so that its equation is

(292) $$x^2+y^2+z^2=(R+\delta)^2\approx R^2+2R\delta,\quad (\delta<<R) \quad ,$$

where $\delta=\delta(P,Q_2)=\delta(\vec{\rho},\vec{h})$ represents the wave aberration in point $P(x,y,z)=P(\vec{\rho},z\approx R)$ relative to the reference sphere in Gaussian image point $Q_2(0,0,0)=Q_2(\vec{h})$.

Note that variable $\vec{\rho}$ determines the point of intersection between emerging rays and the exit pupil plane, while variable $\vec{h}$ represents the measure of the inclination of the central paraxial ray relative to the optical axis. Since this inclination is supposed to be small, the plane xQ_2y basically represents the Gaussian image plane. Furthermore, we will consider Q_2x to be practically parallel to $\vec{h}$ (Fig. 71).

For the following considerations we will take into account the fact that the considered optical system has axial symmetry, so that the aberration of wave $\delta(\vec{h},\vec{\rho})$ generally only depends on the scalar product variables $\vec{h}^2$, $\vec{\rho}^2$, and $\vec{h}\cdot\vec{\rho}$, which are invariant relative to rotation around the optical axis, meaning that $\delta = \delta(h^2,\rho^2,h\rho\cos\theta)$, where θ is the angle between vectors $\vec{\rho}$ and $\vec{h}$. Usually, in order to evince the ***primary geometric aberrations*** (***the Seidel aberrations***), it is sufficient to serially expand the wave aberration δ relative to the rotation invariants up to the second order term, that is,

(293)
$$\delta(h^2,\rho^2,h\rho\cos\theta) = a_0 + b_1h^2 + b_2\rho^2 + b_3h\rho\cos\theta + 1/2[c_{11}h^4 +$$
$$+ c_{12}h^2\rho^2 + c_{13}h^2\cdot h\rho\cos\theta + c_{21}\rho^2\cdot h^2 + c_{22}\rho^4 +$$
$$+ c_{23}\rho^2\cdot h\rho\cos\theta + c_{31}h\rho\cos\theta\cdot h^2 + c_{32}h\rho\cos\theta\cdot \rho 2 +$$
$$+ c_{33}(h\rho\cos\theta)^2] + \dots,$$

followed by the grouping together of the coefficients of identical terms. These coefficients are constants whose value depends on the structure of the optical system considered and the position of the object plane.

Any emergent light ray PP_0 would satisfy the equation of the real wavefront normal, namely

(294)
$$\frac{x-x_0}{\partial F/\partial x} = \frac{y-y_0}{\partial F/\partial y} = \frac{z-z_0}{\partial F/\partial z},$$

in which, according to equation (292),

(295)
$$F(x,y,z) = x^2 + y^2 + z^2 - R^2 - 2R\delta = 0 \quad .$$

Keeping in mind that for a given $Q_2(\vec{h})$, $\delta = \delta(\vec{\rho}) = \delta(x,y)$, it follows from equation (294) that

(296)
$$\frac{x-x_0}{x-R\frac{\partial\delta}{\partial x}}=\frac{y-y_0}{y-R\frac{\partial\delta}{\partial y}}=\frac{z-z_0}{z} \quad ,$$

or, in other words, that

(297)
$$\begin{cases} x-x_0=(1-\frac{z_0}{z})\left(x-R\frac{\partial\delta}{\partial x}\right), \\ y-y_0=(1-\frac{z_0}{z})\left(y-R\frac{\partial\delta}{\partial y}\right). \end{cases}$$

In the following lines we will consider the displacement (the defocusing) z_0 of the Gaussian image relative to plane $z=0$ to be very small, so that we may ignore the product $z_0\cdot\delta$. We will also make the approximation $z\approx R$ in all instances, since the distances along the optical axis to the wavefront at the location of the exit pupil are virtually equal to the radius of the reference sphere.

Under these conditions, the equations at (297) transform into

(298)
$$\begin{cases} x_0=\frac{z_0}{R}x+R\frac{\partial\delta}{\partial x}, \\ y_0=\frac{z_0}{R}y+R\frac{\partial\delta}{\partial y}. \end{cases}$$

Obviously, within the plane of the Gaussian image ($z_0=0$), we have

(299)
$$\begin{cases} x_0=R\frac{\partial\delta}{\partial x}, \\ y_0=R\frac{\partial\delta}{\partial y}, \end{cases}$$

which constitute the components of the vector of ray aberration, $\overrightarrow{Q_2Q_2^*}$.

Because the expression of δ, equation (293), is given in polar coordinates, we will apply the transformations

(300)
$$\begin{cases} x=\rho\cos\theta, \\ y=\rho\sin\theta, \end{cases} \qquad \begin{cases} \rho=(x^2+y^2)^{1/2}, \\ tg\theta=y/x, \end{cases}$$

so that

(301)
$$\begin{cases} \frac{\partial\delta}{\partial x}=\frac{\partial\delta}{\partial\rho}\frac{\partial\rho}{\partial x}+\frac{\partial\delta}{\partial\theta}\frac{\partial\theta}{\partial x}=\cos\theta\frac{\partial\delta}{\partial\rho}-\frac{\sin\theta}{\rho}\frac{\partial\delta}{\partial\theta}, \\ \frac{\partial\delta}{\partial y}=\frac{\partial\delta}{\partial\rho}\frac{\partial\rho}{\partial y}+\frac{\partial\delta}{\partial\theta}\frac{\partial\theta}{\partial y}=\sin\theta\frac{\partial\delta}{\partial\rho}+\frac{\cos\theta}{\rho}\frac{\partial\delta}{\partial\theta}, \end{cases}$$

and the relations at (298) can be written as

(302)
$$\begin{cases} x_0 = \frac{z_0}{R}\rho\cos\theta + R\left(\cos\theta\frac{\partial\delta}{\partial\rho} - \frac{\sin\theta}{\rho}\frac{\partial\delta}{\partial\theta}\right), \\ y_0 = \frac{z_0}{R}\rho\sin\theta + R\left(\sin\theta\frac{\partial\delta}{\partial\rho} + \frac{\cos\theta}{\rho}\frac{\partial\delta}{\partial\theta}\right). \end{cases}$$

By differentiating, and grouping the constants together, from the equations at (293), (302), we obtain the final results

(303)
$$\begin{cases} x_0 = \frac{z_0}{R}\rho\cos\theta + 2Rb_2\rho\cos\theta + Rb_3h + B\rho^3\cos\theta - Fh\rho^2(2+\cos 2\theta) + \\ \qquad +(2C+D)h^2\rho\cos\theta - Eh^3, \\ y_0 = \frac{z_0}{R}\rho\sin\theta + 2Rb_2\rho\sin\theta + B\rho^3\sin\theta - Fh\rho^2\sin 2\theta + Dh^2\rho\sin\theta. \end{cases}$$

in which we've made the replacements B = $2\mathrm{R}c_{22}$, F = $-\mathrm{R}(c_{23}+c_{32})/2$, C = $\mathrm{R}c_{33}/2$, D=$\mathrm{R}(c_{12}+c_{21})$, E = $-\mathrm{R}(c_{12}+c_{31})/2$.
Note that the z_0 terms have been obtained based on the arbitrary defocusing introduced by us relative to the Gaussian image plane (z_0=0).

Let us first discuss the significance of the first order ρ and h terms, which appears because of the b coefficient. Thus, for the b_2 coefficient, in plane $z_0 = 0$ we have

$$\begin{cases} x_0 = 2Rb_2\rho\cos\theta, \\ y_0 = 2Rb_2\rho\sin\theta, \end{cases}$$

so we obtain the circle

$$x_0^2 + y_0^2 = (2Rb_2\rho)^2.$$

In order for this circular spot of light to be reduced to a point, we translate the ideal plane so that $x_0 = y_0 = 0$, that is, we move it to the position $z_0 = -2b_2R^2$, which can be interpreted as either a correction of the longitudinal focalization error, or, if b_2 is dependent on λ, as an ***axial chromatic aberration***. As for the b_3 coefficient, it signifies the transverse translation of the focalizing point within plane $z_0 = 0$ to $x_0 = Rb_3h$, and can also be interpreted either as a correction of the lateral focalization error, or, if b_3 is dependent on λ, as a ***transverse chromatic aberration***.

The geometric aberrations appear in the equations at (303), expressed through the third order terms of ρ and h, which are identified by the ***Seidel coefficients*** B, F, C, D, E (in traditional notation). In order to facilitate classification, we will analyze the

individual contribution of each term while ignoring the contribution of the rest. By doing so, we will obtain ***five types of third order aberrations***, namely: the ***spherical aberration*** ($B \neq 0$), ***coma aberration*** ($F \neq 0$), ***astigmatism*** ($C \neq 0$), ***field curvature*** ($D \neq 0$), and ***distortion*** ($E \neq 0$).

The ***spherical aberration*** ($B \neq 0$), is the only third order aberration that exists along the optical axis ($h = 0$). Let us first consider the ideal plane $z_0 = 0$, in which case,

$$\begin{cases} x_0 = B\rho^3 \cos\theta, \\ y_0 = B\rho^3 \sin\theta, \end{cases}$$

from which, by eliminating θ, we get

(304) $$x_0^2 + y_0^2 = (B\rho^3)^2 \quad .$$

Therefore, the aberration curve is a circle with its center in the paraxial image point Q_2 and a radius equal to $B\rho^3$, and constitutes the ***transverse spherical aberration*** (Fig. 72). The image of any object point is therefore a circular diffusion spot whose radius is proportional to the cube of the exit pupil radius. Obviously, this aberration does not discomfort the eye if its diameter of $2B\rho^3$ corresponds to a smaller viewing angle than the angular magnification $(\gamma_1)_{min}^{eye} \approx 1'$.

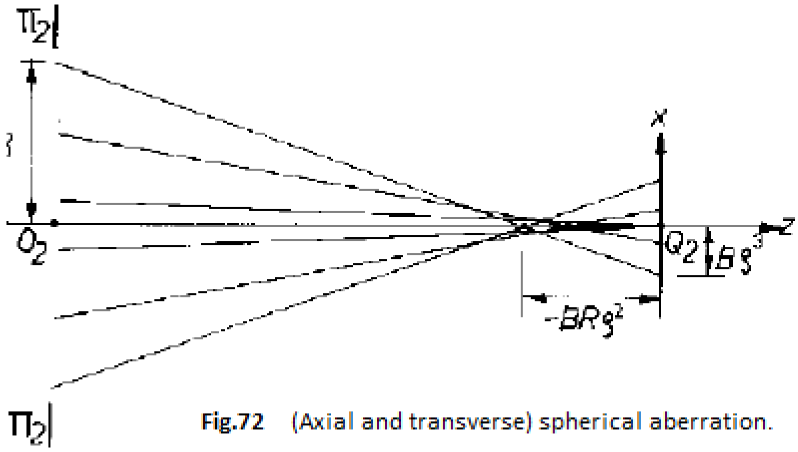


Fig.72 (Axial and transverse) spherical aberration.

Any ray exiting at a distance ρ from the center of the exit pupil will generally intersect the O_2z axis in a point other than the Q_2 point of paraxial ray convergence. If we refer to the condition for intersection, $x_0 = y_0 = 0$, within a plance $z_0 \neq 0$, based on the equations at (303), we obtain

$$\begin{cases} x_0 = \frac{z_0}{R}\rho\cos\theta + B\rho^3\cos\theta = 0, \\ y_0 = \frac{z_0}{R}\rho\sin\theta + B\rho^3\sin\theta = 0, \end{cases}$$

which give the ***axial spherical aberration*** (see Fig. 72)

(305) $$z_0 = -BR\rho^2$$

The sign of this aberration can be positive or negative. Thus, for example, in the case of a convergent lens, $z_0 < 0$, and in that of a divergent lens, $z_0 > 0$, from which fact we may deduce the possibility of reducing the spherical aberration by combining these two types of lens. Usually, according to equations (304), (305), the spherical aberration only depends on the ρ variable, which constitutes a measure of the angular aperture γ, hence the term ***aperture aberration***.

All the other geometric aberrations indicated by equations (303) appear because of light beams that are skewed relative to the optical axis ($h \neq 0$). They therefore only affect the images of extra-axial points, and are of a first, second, or third order relative to the h variable, which is a measure of the viewing angular field θ. We may gather these aberrations caused by skewed light beams under the umbrella term ***field aberrations***. From among these we will first consider the ***coma aberration*** ($F \neq 0$). In this case, within the ideal plane $z_0 = 0$, we have

$$x_0 = -Fh\rho^2(2+\cos 2\theta), \qquad y_0 = -Fh\rho^2 \sin 2\theta,$$

based on which, by eliminating the angle 2θ, we obtain the aberration curve equation

(306) $$(x_0 + 2Fh\rho^2)^2 + y_0^2 = (Fh\rho^2)^2 \quad ,$$

which is an equation for a circle with its center in $(-2Fh\rho^2, 0)$ and a radius of $Fh\rho^2$. Consequently, because of the chromatic aberration, the light rays exiting through the exit pupil ring of radius ρ form a circle within the Gaussian image plane, above or below point Q_2, depending on the sign of coefficient F (Fig. 73). The overlaying of these circles, corresponding to the entire surface of the exit pupil, constitute the image of an object point within said plane. The image thus obtained has an elongated shape, similar to the tail of a comet, hence the term coma.

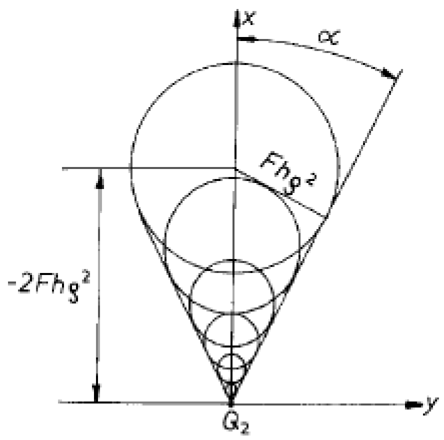


Fig.73. Coma.

Note that, because of the dependence of coordinates x_0 and y_0 on the angle 2θ, a single rotation along the circle $\rho = const.$ from within the exit pupil plane leads to a double rotation along the circle in the Gaussian image plane. Moreover, the envelopes of circles from within the Gaussian image plane are two line segments that intersect each other in the paraxial image point Q_2, under an angle of 2α, where

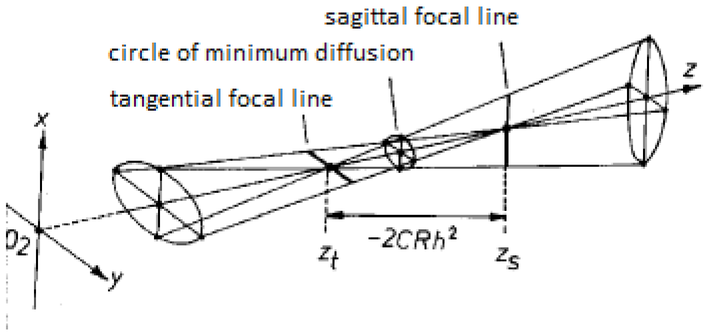


Fig. 74 Astigmatism and the circle of minimum diffusion.

$$\alpha = \arcsin\left(Fh\rho^2/2Fh\rho^2\right) = \arcsin(1/2) = 30^0$$

as can be seen in Fig. 73.

Unlike coma aberrations, in whose case the images of points extend across a plane perpendicular to the optical axis (Fig. 73), astigmatism causes them to extend along it (Fig. 74).

Next we will discuss ***astigmatism*** ($C \neq 0$) and ***field curvature*** ($D \neq 0$), which take place in combination. If we consider these aberrations within a plane of ray interception $z_0 = const.$, based on equations (303), we have

$$\begin{cases} x_0 = \frac{z_0}{R}\rho\cos\theta + (2C+D)h^2\rho\cos\theta, \\ y_0 = \frac{z_0}{R}\rho\sin\theta + Dh^2\rho\sin\theta, \end{cases}$$

from which, if we eliminate the angle θ, we obtain the equation for the aberration curve

(307) $$\frac{x_0^2}{\left[\frac{z_0}{R}\rho + (2C + Dh^2\rho)\right]^2} + \frac{y_0^2}{\left[\frac{z_0}{R}\rho + Dh^2\rho\right]^2} = 1.$$

The image of an extra-axial object point Q_1 thus appears as a spot of light bordered by the ellipsis given by equation (307), centered along the axis O_2z and with the axes parallel to the coordinate axes O_2x and O_2y (Fig. 74). If we translate the plane $z_0 = const.$, the image remains elliptical, but its shape and dimensions change. For

two plane positions, the ellipsis degenerates into linear segments (focal lines), one of which is parallel to O_2x, and the other to O_2y. By eliminating the ellipsis semi-axes in equation (307), we obtain the position $z_0 = z_s$ of the ***sagittal focal line*** plane, formed by the equatorial rays, and the position $z_0 = z_t$ of the ***tangential focal line*** plane, formed by the meridional rays*, namely

(308) $$z_s = -(2C + D)Rh^2, \qquad z_t = -DRh^2.$$

The difference $z_s - z_t = -2CRh^2$ constitutes a measure of the astigmatism (C) for the object point Q_1 considered (with given R, h). For $z_0 = (z_s + z_t)/2$, we obtain the highest degree of light ray concentration within a circular disk, termed ***circle of minimum diffusion*** (confusion), or pseudo-focus (Fig. 74).

Let us next consider a straight line normal to the optical axis object point P_1. Every point along this line has a corresponding pair of focal lines as image. Through rotation around the optical axis, we obtain a transverse plane object and its image, a curved surface, the envelope of focal lines, termed ***caustic***. According to the equations at (308), this surface has two sheets, namely the ***sagittal field*** (the geometrical location of positions z_s) and the ***tangential field*** (the geometrical location of positions z_t), which have the shape of rotational paraboloids with a common vertex in the paraxial image point P_2 (Fig. 75). This deviation from the Gaussian image plane is termed ***field curvature***, or ***Petzval curvature***.

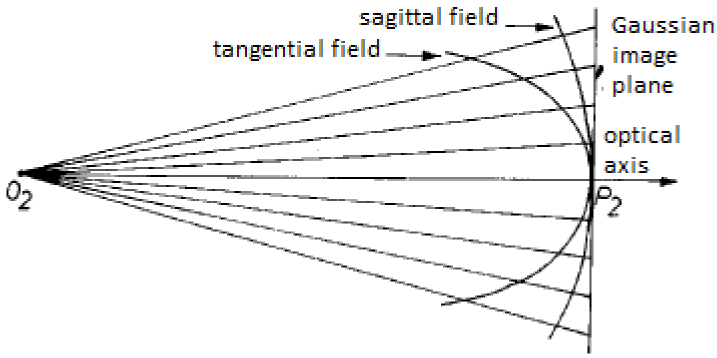


Fig. 75 Field (Petzval) curvature.

The last third order aberration is ***distortion*** ($E \neq 0$). According to the equations at (303), within the Gaussian image plane $z_0 = 0$ we have

$$x_0 = -Eh^3, \quad y_0 = 0\,.$$

* The meridional plane is defined by the object point Q_1 and the optical axis.

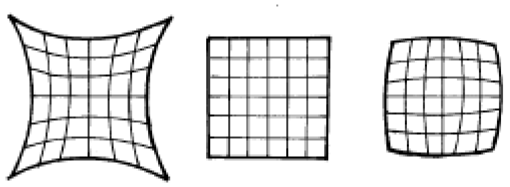


Fig. 76 Distortion.

Since it is independent of ρ and θ, distortion no longer determines the spread of light around the position of the ideal image Q_2, but only its transverse translation to another point by a value proportional to h^3. The image of a spatially extended object remains clear, but deformed. The only exception is the lines from within the object plane that intersect the optical axis. All the other lines have curves as images, specifically convex curves if $E < 0$ and concave curves if $E > 0$, with respect to the optical axis. These "pillow" and "barrel" shaped distortions are easily highlighted using a simple figure, of finite spatial extension, such as the rectangular web (Fig. 76). Distortion is troublesome if precise extra-axial measurements are required.

The primary geometric aberrations described above appear when light rays leave the paraxial domain, and are caused by the finite value of the aperture (the variable ρ) and/or of the view field (the variable h). The proportions between these aberrations are given by the Seidel coefficients, that, for a given objet position, depend on the shape, thickness, the distances between, and the refractive indices of the optical system components, and the position of the diaphragms. Explicitly and analytically determining the Seidel coefficients relative to the multitude of parameters on which they depend is an extremely difficult task, even for simple optical systems.

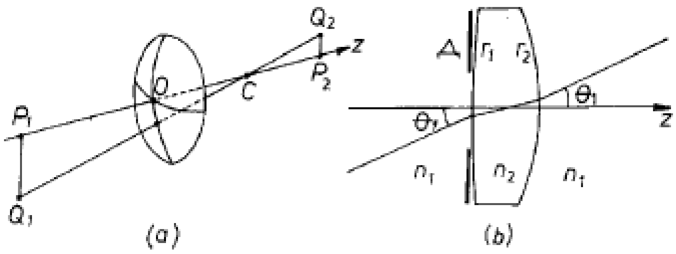


Fig. 77 Calculating the astigmatism of thin lenses.

In order to illustrate, let us consider a diaphragmed thin lens (Fig. 77, b), so that it is traversed by light rays only in the vicinity of the optical axis Oz, where the lens behaves almost like a plate with planar, parallel faces, and so the angles of incidence and emergence are practically equal (θ_1).

Let us calculate the astigmatism for this lens by twice applying the Young formulae for spherical diopters (see Chapter 2.1, equations (115) and (116)), in this instance the light beam originating from an extra-axial object point Q_1 (Fig. 77a). Thus, if we consider the extreme immersion media to be identical ($n_1 = n_3$) and if we write the relative refractive index of the lens as $n = n_1/n_2$, for a fan of meridional rays

(within the plane Q_1Oz), equation (115) leads to

$$(309)\qquad \begin{cases} \dfrac{\cos^2\theta_1}{l} + \dfrac{n\cos^2\theta_2}{s_2^{\prime}} = \dfrac{n\cos\theta_2 - \cos\theta_1}{r_1}, \\ \\ -\dfrac{n\cos^2\theta_2}{s_2^{\prime}} + \dfrac{\cos^2\theta_1}{t} = -\dfrac{n\cos\theta_2 - \cos\theta_1}{r_2}, \end{cases}$$

which added together give the t position of the tangential field:

$$(310)\qquad \frac{1}{l} + \frac{1}{t} = \frac{n\cos\theta_2 - \cos\theta_1}{\cos^2\theta_1}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad .$$

Similarly, for a fan of equatorial (sagittal) rays, equation (117) leads us to

$$(311)\qquad \begin{cases} \dfrac{1}{l} + \dfrac{n}{s_2^{\prime\prime}} = \dfrac{n\cos\theta_2 - \cos\theta_1}{r_1}, \\ \\ -\dfrac{n}{s_2^{\prime\prime}} + \dfrac{1}{s} = -\dfrac{n\cos\theta_2 - \cos\theta_1}{r_2}, \end{cases}$$

which added together give the s position of the sagittal field:

$$(312)\qquad \frac{1}{l} + \frac{1}{s} = (n\cos\theta_2 - \cos\theta_1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad .$$

Particularly, for $\theta_1 = \theta_2 = 0$, the oblique abscises l, s, t become regular abscises (oriented along the optical axis), namely l_0, s_0, t_0, so that equations (310) and (312) lead us to the thin lens formula

$$(313)\qquad \frac{1}{l_0} + \frac{1}{t_0} = \frac{1}{l_0} + \frac{1}{s_0} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = \frac{1}{f} \quad .$$

Next we will consider a spherical object surface AOB of radius $OC = \rho$, situated symmetrically relative to the optical axis Oz (Fig. 78). If we use the segment notations $AP = l$, $OP = l_0$, and $\angle APO = \theta_1$, we have

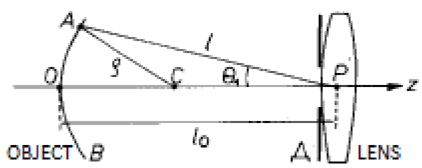


Fig. 78 Calculating curvatures $1/\rho_s$ and $1/\rho_t$, and deducing the Petzval condition.

$$\rho^2 = l^2 + (l_0 - \rho)^2 - 2l\,(l_0 - \rho)\cos\theta_1 \quad ,$$

or, if we approximate $\cos\theta_1$ to $1 - (\theta_1^2/2)$,

$$\rho^2 = (l - l_0 + \rho)^2 + l\,(l_0 - \rho)\theta_1^2$$

If we serially expand the radical of the second member and take the first two terms separately, we obtain the following expression:

$$\rho = l - l_0 + \rho + \frac{1}{2}\frac{l\,(l_0 - \rho)}{l - l_0 + \rho}\theta_1^2 \qquad .$$

Rearranging the terms, we get:

(314) $$\frac{1}{l} = \frac{1}{l_0} + \frac{1}{2}\left(\frac{1}{\rho} - \frac{1}{l_0}\right)\theta_1^2 \qquad ,$$

which is valid in the vicinity of the optical axis. We will use this formula for the curvature $1/\rho$ of the object surface, as well as for the curvatures $1/\rho_t$ and $1/\rho_s$ of the tangential and sagittal fields, respectively, that is,

(315) $$\begin{cases} \frac{1}{t} = \frac{1}{t_0} + \frac{1}{2}\left(\frac{1}{\rho_t} - \frac{1}{t_0}\right)\theta_1^2, \\ \frac{1}{s} = \frac{1}{s_0} + \frac{1}{2}\left(\frac{1}{\rho_s} - \frac{1}{s_0}\right)\theta_1^2, \end{cases}$$

in which the same angle θ_1 is used, since the angle of emergence from the lens is virtually equal in value to the angle of incidence (Fig. 77b). If we next introduce expressions (314) and (315) into equations (310) and (312), and take into account the relations at (313) and the approximations $\theta_1 = n\theta_2$, $\cos\theta \approx 1 - (\theta^2/2)$, and $1/\cos\,^2\theta \approx 1 + \theta^2$, we obtain the simple result

(316) $$\begin{cases} \frac{1}{t} = \frac{1}{t_0} + \frac{1}{2}\left(\frac{1}{\rho_t} - \frac{1}{t_0}\right)\theta_1^2, \\ \frac{1}{s} = \frac{1}{s_0} + \frac{1}{2}\left(\frac{1}{\rho_s} - \frac{1}{s_0}\right)\theta_1^2, \end{cases}$$

Particularly, for a plane object surface ($\rho \to \infty$), we have

(317) $$\begin{cases} \frac{1}{\rho_t} = \frac{3n+1}{nf}, \\ \frac{1}{\rho_s} = \frac{n+1}{nf}. \end{cases}$$

Notice that for any given lens, the field curvatures have the same orientation, the curvature $1/\rho_t$ of the tangential field being higher than the curvature $1/\rho_s$ of the sagittal field (also see Fig. 75). The field curvatures for the convergent lens ($1/f > 0$) and the divergent lens ($1/f < 0$) are opposite in orientation. There is thus the possibility of compensating for the astigmatism and field curvature, by designing

centered systems of diaphragmed thin lenses attached together. In this case, the field curvatures are algebraically summed up as

(318)
$$\begin{cases} \frac{1}{\rho_t} = \sum_i \frac{3n_i+1}{n_i f_i} = \frac{3}{f} + \sum_i \frac{1}{n_i f_i}, \\ \frac{1}{\rho_s} = \sum_i \frac{n_i+1}{n_i f_i} = \frac{1}{f} + \sum_i \frac{1}{n_i f_i}, \end{cases}$$

where we've used the notation $1/f = \sum_i 1/f_i$ for the system convergence.

The condition $\rho_t = \rho_s$ for eliminating astigmatism therefore implies the condition $1/f = 0$, meaning the system must be afocal (and therefore equivalent to a plane-parallel plate), in which case the equations at (318) become

(319)
$$\frac{1}{\rho_t} = \frac{1}{\rho_s} = \sum_i \frac{1}{n_i f_i} \quad .$$

If we also require the elimination of the field curvature, in which case $1/\rho_t = 1/\rho_s = 0$, we obtain the ***Petzval condition***,

(320)
$$\sum_i \frac{1}{n_i f_i} = 0 \quad .$$

Note that the aberration can also be softened in systems of finite focal distance ($1/f \neq 0$) as long as they satisfy the Petzval condition (320). In this case, $\rho_s = 3\rho_t = f$, as can be deduced from the equations at (318).

Usually field curvature of a certain degree is tolerated in visual instruments, because the eye has the ability to adjust to it. The requirements for the objectives of photographic or projection apparatuses are however much higher. If meeting the Petzval condition is not enough, or if it cannot be met, the field can be significantly corrected within the axial domain by the use of a ***field flattener*** positioned in the immediate vicinity of the image plane. The curvature of such a lens is usually small so as to not induce aberrations.

Here are several other results concerning thin lens aberrations (with demonstrations omitted):

In the case of objects located far away, the curvature radii at which spherical aberration is minimized are:*

* See V.V. Bianu, *Optica geometrică*, Ed. Tehnică (Technical Press), Bucharest, 1962 (§ 126, pp 208 – 215).

(321)
$$\begin{cases} r_1 = \frac{2(n+2)(n+1)}{n(2n-1)} f, \\ r_2 = -\frac{2(n+2)(n-1)}{4+n-2n^2} f, \end{cases}$$

and those at which coma aberration is minimized are:

(322)
$$\begin{cases} r_1 = \frac{n^2-1}{n^2} f, \\ r_2 = \frac{n^2-1}{n^2-n-1} f. \end{cases}$$

Lens designers are confronted with the necessity of an optimal compromise between criteria (321) and (322).

In order to minimize chromatic and geometrical aberrations and find an optimum configuration relative to the use intended, the designer of optical systems must manipulate a multidimensional multitude of variables (refractive index, geometrical shape, thickness, distance, diaphragm, etc.). We've presented a simple example in paragraph 2.7, where we showed how chromatic aberrations can be compensated for by combining lenses of different types of glass. Said lenses display inverse spherical aberrations that depend of their shapes. They also display inverse coma aberrations that depend on the lens shapes as well, but in a different fashion. That is why the degrees of freedom available when achromatizing the system are often used as much as possible to also compensate for spherical and coma aberrations.

There are also widely used methods by which the third order aberrations previously discussed are partly counterbalanced by manipulating corresponding higher order aberrations. For such refined adjustments, as well as for the automatic optimization of optical system parameters in general, elaborate calculation software has been developed for the ***precise calculation and depiction of the paths of light rays*** passing through given systems and for the ***generation of spot diagrams***, displaying the points of intersection between rays and various given planes perpendicular to the optical axis. That software allows users to start from a regular polygon of points (usually a square) within the entrance pupil plane and see how the densities of points of intersection with any ulterior plane (such as the Gaussian image plane or the planes in its proximity) are distributed. This constitutes a direct measure of the distribution of the luminous flux in that plane.

But starting with aberration classical theory, founded by Hamilton (with his characteristic function) and Bruns (with his eikonal function), and continuing to the present day, analytical methods of studying the geometrical properties of wavefronts

have been and are still under intense investigation.

Unlike planar, spherical, or cylindrical waves, which arise occasionally, common wavefronts are generally incredibly complicated. Although they appear even and approachable enough the further they are from the focalized area, wavefronts considered nearer this space become more and more difficult to visualize and almost impossible to analyze. The ***Twyman-Green interferometer*** (Fig. 79) is a very powerful device for precise optical measurements, especially of wavefronts generated by the optical components. It is a version of the Michelson interferometer, and consists of a point source S of monochromatic light positioned at the focus of a lens L_1 and a perfectly spherical mirror O_2 with its center C positioned at the Gaussian focus of the lens or optical system to be tested. If said system is free of aberrations, the wave reflected in O_2 back towards the beam splitter D will be perfectly planar, and the field of interference with the planar wave reflected in mirror O_1 will appear uniform. If, however, aberrations deform the wavefront passing through the analyzed system L and back, then the deformity will be clearly indicated through the interference fringes (lines of equal phase difference from the reference planar wave). The interference image can be viewed with the naked eye or can be photographed. Fig. 79 illustrates the Twyman-Green interferograms associated with spherical (a) and coma (b) aberrations, corresponding to the paraxial focal plane, and with astigmatism, corresponding to the minimal diffusion circle plane (c), to the plane of a focal line (d), or to some other plane (e). Any local variation of optical path, even of the magnitude of wavelength fractions, caused by imperfections of the system surfaces or by the existence of spaces of non-homogeneous refractive index, lead to the deformation of the wavefront and generate corresponding fringes in the interference field, which allow those irregularities to be localized. The method discussed affords the significant advantage of the complete wavefront shape being displayed immediately, as the

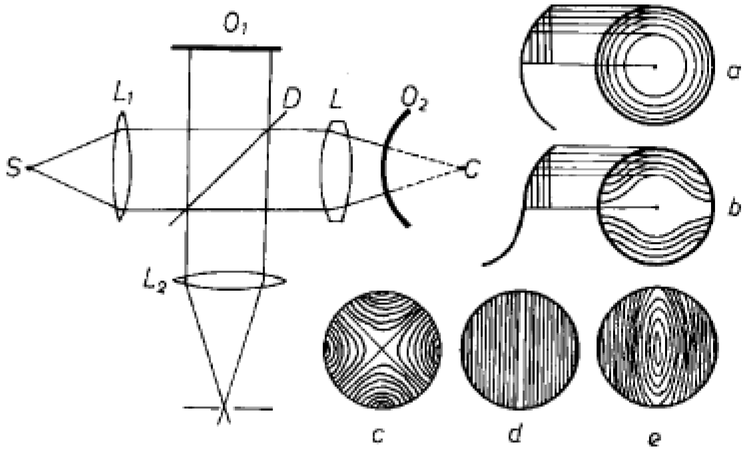


Fig. 79 The Twyman-Green interferometer and several interferograms.

structures displayed to the left of the meridional plane interferograms (a) and (b) illustrate. As is usually the case, the performance of the interferometer significantly increases if the conventional monochromatic light source is replaced with a laser light source.

We've constantly discussed lenses and optical systems involving spherical surfaces and the geometrical aberrations inherent only to them. Due to the relative facility and high precision with which spherical surfaces are fashioned, their use has spread most widely, in contrast to the Cartesian surfaces (see Chapter 1.3). Still, stigmatic aspherical surfaces and elements are frequently incorporated into high performance image forming optical systems.

After precisely identifying the deformities of wavefronts generated by a real optical system using the interferometer method described above, the question arises of how to make the suitable adjustments. In this last segment we will therefore present a short description of the important practical issue of ***correcting surfaces***. We will show how any astigmatism-related flaws (concerning a given pair of conjugate points) produced upon the passage of an initially homocentric beam through an optical system can be eliminated by making the appropriate corrections to the surface of separation of the last refringent diopter of the system. To do so, we will refer to Huygens' construction (Fig. 10 of Chapter 1.2) and reformulate the problem as follows: given the (deformed) wavefront ϕ_1 passing through the second to last medium (n_1) of an optical system, let us determine the shape of the last separation surface Σ so that the wavefront ϕ_2 passing through the final medium (n_2) should be spherical. The position of surface Σ will thus be determined based on the intersections of those lines normal to ϕ_1 and ϕ_2, whose segments s_1, s_2 ending at Σ meet the condition for stigmatism $n_1 s_1 + n_2 s_2 = constant$. Obviously, the infinity of values that can be given to the constant in this formula allows for the existence of an infinity of possible surfaces Σ.

In practice it is often enough to apply ***local optical corrections***. Let us demonstrate by considering the common situation of a final refringent surface Σ separating a medium of index n from air (of index n = 1) and of a random wavefront ϕ passing through the final medium that slightly deviates from a theoretical wavefront ϕ_0 corresponding to the ideal condition of rigorous stigmatism

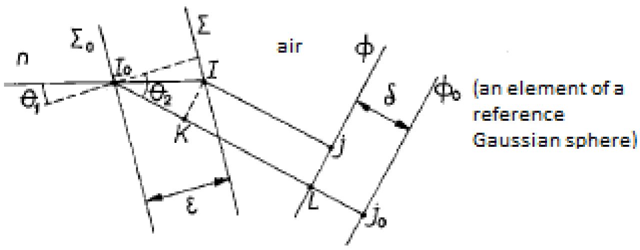


Fig. 80 The principle of optical

(Fig. 80). In order to compensate for a local aberration δ, it thus suffices to slightly polish the initial surface Σ so that it matches the correct shape Σ_0. The relation between the local depth to which the surface must be polished and the wavefront aberration δ that must be eliminated is easy to determine. The rule for making optical corrections is written as follows:

(323) $$nI_0I + IJ = I_0J_0 \quad ,$$

or, if we take into consideration the fact that $IJ = KL$, as:

(324) $$nI_0I = I_0K + \delta \quad .$$

But by studying the geometry of Fig. 80, we readily deduce that:

(325) $$I_0I = \varepsilon/\cos\theta_1 \quad , \quad I_0K = I_0I \cdot \cos(\theta_2 - \theta_1) \quad .$$

From these three equations we obtain the sought for relation between ε and δ, namely:

(326) $$\varepsilon = \frac{\cos\theta_1}{n - \cos(\theta_2 - \theta_1)} \cdot \delta \approx \frac{\delta}{n-1} \quad .$$

Similar calculations concerning a reflecting surface leads us to the formula:

(327) $$\varepsilon = \frac{\delta}{2\cos\theta_1} \approx \frac{\delta}{2} \quad .$$

Since the paths followed by the various rays connecting rigorously stigmatic conjugate points must contain the same number of wavelengths λ, it follows that the tolerance correlated with the aberration δ, meaning the tolerance values of ε are of the magnitude of small wavelength fractions.

Chapter III
NON-HOMOGENEOUS MEDIA

The continuous curvature of light rays passing through non-homogeneous media explains various phenomena such as those of ***atmospheric refraction***. Because the density, and therefore the refractive index, of air decreases with altitude, light rays coming from stars curve across a path concave towards the Earth (see Chapter 1.1, equation (16)). For this reason, the apparent height of a given star relative to the horizon is greater than its actual height (***regular astronomical refraction***). ***Irregular refraction***, caused by atmospheric turbulence, produce ***stellar scintillation***. Similar effects, termed ***mirages***, occur near the Earth's or other surfaces when their temperature is greater or smaller than that of the surrounding air, thus inducing a temperature, and so a density and a refractive index, gradient.

The calculation of trajectories of light rays passing through continuous non-homogeneous media is generally considered important in many applications. As a first example let us consider the ***gas lens***, generally comprising of a heated cylindrical metallic tube through which a laminar flow of gas passes. The concentration of gas, and so its refractive index, is greater along the cylinder's axis, where the temperature is lower. For this reason, the light rays transmitted through the tube are deviated towards said axis, an effect similar to that of a lens. Compared to a glass lens, the gas version completely avoids losses of light caused by reflection at the surfaces separating air from glass and by the diffusion suffered across the surface of dust particles or of imperfections across the lens surface. Of particular importance are the gradated planar or cylindrical and symmetrical structures used in designing ***optical guides*** (optical plates and fibers).

A high precision method of visualizing irregularities of the refractive index, based on the way they divert light rays, was advanced by A. Töpler (1864), and is known as the ***schlieren technique***, or ***streak method*** (from the German word "die schliere", which translates into "stria"). Its principle is illustrated in Fig. 81.a, in which the image of (point or slit) light source S is formed across a small opaque disc (or "knife-edge") D, using lenses L_1, L_2, between which the schlieren camera (C),

containing the transparent optical medium to be investigated, is mounted. A lens L_3 is used to form the images of the various planes of this medium across a screen (E). As can be surmised, if the medium is homogeneous, the screen will appear darkened due to the disc's acting liking a screen. If, however, there are irregularities within the medium, the deflected light rays will go around the disc and form a ***schlieren image*** across the screen, that is, a map of the local refractive index gradients causing their being deviated (and focused on the screen by way of lens L_3). A method complementing the aforementioned technique is that of the ***shadowgraph***, which replaces the disc used in the previous method with an opaque screen having a small aperture in its center. In this case, the screen will only be illuminated by the light rays not being diverted, since the diverted rays will be blocked by the screen and excluded from the initial light beam. The screen will display a ***shadowgram***, in which the darkened lines corresponding to the points within the investigated medium of irregular refractive index will appear against a light background.

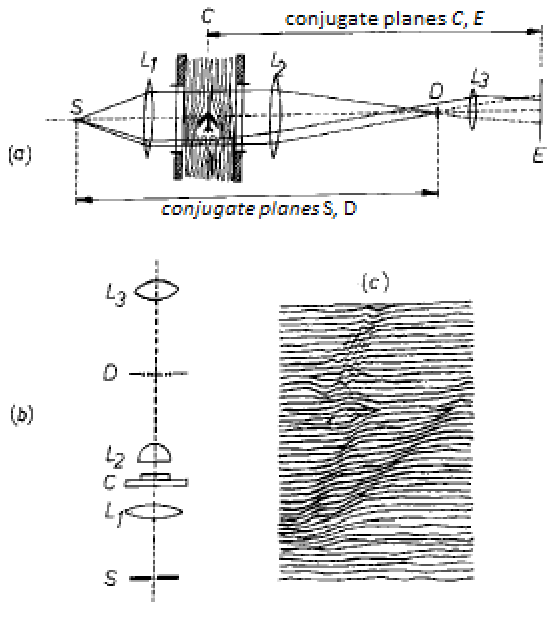


Fig. 81 Principle of the schlieren method

Numerous imaginative advancements have also been made within this field. An example would be the use of the schlieren method in combination with microscopes (J.R. Meyer–Arendt, 1961), as illustrated in Fig. 81.b. A one-dimensional cutoff pattern (D) of 5 – 10 lines/mm parallel to the slit S is mounted between the objective (L_2) and the ocular (L_3). If the object under investigation, placed on the glass slide (C), is homogeneous, its image will appear furrowed by equidistant lines corresponding to the pattern. Any variation in the refractive index causes specific distortions, as illustrated in Fig. 81.c, which describes the phenomenon when the glass slide is partially covered with a transparent substance.

Another continuous non-homogeneous "optical" medium is obtained in ***electronic and ionic optical*** devices, particularly in electronic and ionic microscopes. The non-relativistic motion of particles is described by the laws of classical mechanics, and, as we've seen in Chapter 1.2, in the case of conservative force fields, satisfies the

Maupertuis-Euler principle, analogous to the Fermat principle. In these cases, particle speed fulfils the role of refractive index n. For example, if we consider the motion of charged particles within an electrostatic field and choose a convenient zero value for electric potential U, we may consider the analogy $n = \sqrt{U}$, where U must satisfy the ***Laplace equation*** $\Delta U = 0$. If this condition is satisfied, all the light ray equations of geometrical optics become valid for the geometrical optics of electron and ion trajectories. Particle trajectories are normal to the group of surfaces satisfying the ***eikonal equation*** $|\nabla\phi| = \sqrt{U}$, the ***expansion*** of the particle beam, $Ud\Omega dS\cos\gamma$, is invariable during propagation, the radiance of the object and that of the image satisfy the ***Clausius theorem*** $L_1/U_1 = L_2/U_2$, and so on.

Working with geometrical optics approximations, a general calculating method for optical fields entails determining wavefronts $\phi(\vec{r}) = constant$, by integrating the eikonal equation $(\nabla\phi)^2 = n^2$ and tracing the light rays with the use of the equation $\nabla\phi = n\vec{\tau}$. Alternatively, the light ray and (canonical) Hamilton or Euler-Lagrange equations may be integrated directly.

We will next illustrate one last method, appropriate for simple situations in which the distribution of the refractive index exhibits certain symmetries, namely translational (planar structures), cylindrical, and spherical symmetries.

3.1 Planar Structures

Let us consider the one-dimensional function of the refractive index, of general form $n = n(x)$, within a Cartesian system of coordinates. Advancement along the trajectory is thus given by the equation:

(328) $$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2 \quad .$$

As the refractive index is only dependent on variable x, it will be convenient to consider the light ray equation relative to the other two components, as:

(329) $$\frac{d}{ds}(n\tau_y) = 0 \quad , \qquad \frac{d}{ds}(n\tau_z) = 0 \quad ,$$

which would mean that the values of $n\tau_y$ and $n\tau_z$ are conserved along the trajectory, that is,

(330) $$ndy/ds = A \quad , \qquad ndz/ds = B \quad ,$$

in which A and B are constants determined based on initial conditions. From equations (328) and (330) we also obtain the relation:

(331) $$n dx/ds = \sqrt{n^2 - (A^2 + B^2)} \quad .$$

Parameter s is eliminated through divisions between the equations at (330) and (331), following which we obtain the following system:

(332) $$\begin{cases} dy/dz = A/B, \\ dy/dx = A/\sqrt{n^2(x) - (A^2 + B^2)}, \\ dz/dx = B/\sqrt{n^2(x) - (A^2 + B^2)}, \end{cases}$$

which if we integrate we obtain the projections of the ray trajectory onto the three coordinate planes. It will be observed that the projection of the trajectory onto the yOz plane, irrespective of $n(x)$ distribution, is the line:

(333) $$y = (A/B)z + C \quad .$$

In other words, the trajectory is a planar curve within the plane defined by equation (333). For this reason, without diminishing the general character of these considerations, we will study the trajectory within the plane $y = 0$, that is, we will set the value 0 for constants A and B, as is illustrated in Fig. 82. In this case, constant B of equation (330) is given by the equation:

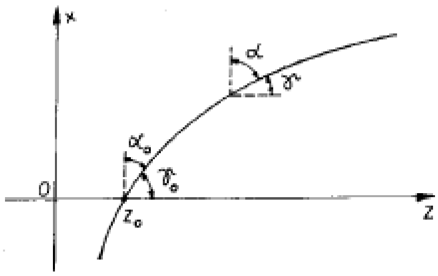


Fig. 82 A light ray trajectory and its local parameters.

(334) $$B = n\tau_z = n dz/ds = n\cos\gamma = n_0\cos\gamma_0 \quad ,$$

in which n_0 and γ_0 correspond to the "launch point" $x = 0$, $z = z_0$. Evidently, equation (334) is the ***Snell-Descartes law*** $n \sin\alpha = n_0 \sin\alpha_0$ applied to media of planar structure. For $A = 0$ and $B = n_0 \cos\gamma_0$, the third integrated equation represents the ray trajectory

(335) $$z(x) - z_0 = \int_0^x \frac{n_0\cos\gamma_0}{\sqrt{n^2(x) - n_0^2\cos^2\gamma_0}} dx \quad .$$

The range of values for x that allow the propagation of the light rays can be identified based on the condition that the integrand be real, that is:

(336) $$n(x) \geq n_0 \cos\gamma_0 \quad .$$

When $n(x) < n_0$ and $dn/dx < 0$ (Fig. 83), equation (336) imposes a limited value $x = x_{max}$ given by the equation:

(337) $$n(x_{\max}) = n_0 \cos\gamma_0 \quad .$$

This liminal value, for which $dx/ds = 0$ (see equation (336)), corresponds to the ***turning point***, or ***point of total reflection***. The wider the initial trajectory angle γ_0, the greater the entry distance x_{max} (Fig. 83) is. Evidently, total reflection does no take place if $n(x) > n_0$ and $dn/dx > 0$.

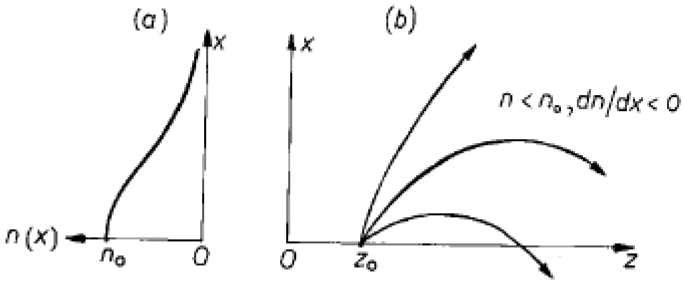


Fig. 83 Group of light trajectories (b) corresponding to distribution (a).

Let us next consider an instance of distribution of planar symmetry, that is, a medium in which the refractive index decreases symmetrically on both sides of value n_0 within plane $x = 0$. In this case, the trajectories are confined between an upper limit x_{max} and a lower limit x_{min}. This sort of distribution has particular practical importance in guiding light through optical circuits. An example would be ***parabolic distribution***, commercially branded ***Selfoc***, for which the refractive index is given by the equation:

(338) $$n^2 = n_0^2\left(1 - \frac{x^2}{a^2}\right) \quad ,$$

in which n_0 and a are constants (Fig. 84). By calculating the integral in the trajectory equation, equation (335), applied to distribution (338), and considering the launch point at the origin $O(x = 0, z = 0)$ of the coordinates, we obtain:

(339) $$z(x) = a\cos\gamma_0 \cdot \arcsin\left(\frac{x}{a\sin\gamma_0}\right) \quad ,$$

or, if we invert the direction of dependence:

(340) $$x(z) = a\sin\gamma_0 \cdot \sin\left(\frac{z}{a\cos\gamma_0}\right) \quad .$$

Equation (340) represents a sine trajectory (Fig. 84) of amplitude $x_{max} = a\sin\gamma_0$ (as can also be directly surmised from the condition for total reflection (337) and distribution (338)) and the spatial half-period $\Delta z = \pi a \cos\gamma_0$. As the initial angle γ_0 increases, the amplitude increases and the period decreases. As expected, based on the general considerations just discussed, light is confined to propagate between limits $\pm x_{max}$. However, the remarkable feature of the Selfoc system is that it focuses all paraxial rays ($\cos\gamma_0 \approx 1$) in the same point after each half-period $\Delta z \approx \pi a$. The focus points O', O'', etc. are veritable successive images of the launch point O, as the optical path between two consecutive focalizations is practically the same in the case of paraxial rays. Indeed, if we consider equations (334), (338), (340), we have:

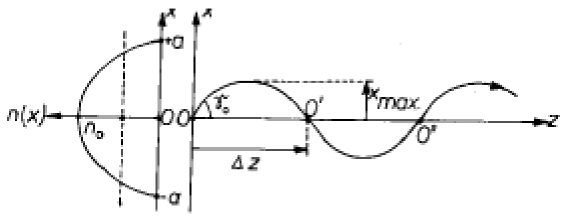


Fig. 84 Sine trajectory in the planar Selfroc structure.

(341) $$[OO'] = \int_O^{O'} n ds = \int_0^{\Delta z} n\frac{ds}{dz}dz = (n_0\cos\gamma_0)^{-1}\int_0^{\Delta z} n^2[x(z)]dz = \pi a n_0 + O(\gamma_0^2) \quad ,$$

where $O(\gamma_0^2)$ represents terms in γ_0 from the second degree onwards. This property may be explained by referring to the fact that, as we consider rays of wider launch angle γ_0, the geometrical path traversed between two points of focalization becomes longer than the axial ray ($\gamma_0 = 0$), but since they mostly pass through an area of smaller refractive index, the optical path remains the same.

3.2 Cylindrical structures

Let us consider a medium whose refractive index only depends on the distance r to a fixed axis Oz. In this case, we will first describe the trajectory through a representation based on $r(z)$, $\theta(z)$, z (see Fig. 13 in Chapter 1.2), in order to illustrate the application of Lagrangian formalism for the cylinder coordinates r, θ, z. The



element describing the trajectory is therefore (see Fig. 85):

(342) $$ds = \sqrt{(dr)^2 + (rd\theta)^2 + (dz)^2} = (1 + r'^2 + r^2\theta'^2)^{1/2} dz \quad ,$$

where we've used the notations $r' = dr/dz$, $\theta' = d\theta/dz$, so that the ***optical Lagrangian*** takes the form (see equation (50) in Chapter 1.2)

(343) $$L(r, \theta, r', \theta', z) = n\frac{ds}{dz} = $$
$$= n(r) \cdot (1 + r'^2 + r^2\theta'^2)^{1/2}$$

Evidently, the expression (343) for the Lagrangian for cylinder coordinates can be deduced from its expression in the case of Cartesian coordinates:

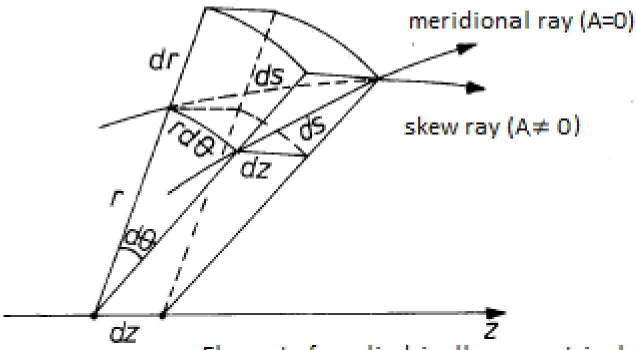


Fig. 85 Element of a cylindrically symmetrical structure.

(344) $$L(x, y, x', y', z) = n\frac{ds}{dz} = n(x, y, z) \cdot (1 + x'^2 + y'^2)^{1/2} \quad ,$$

within which we effect the transformation $x = r\cos\theta$, $y = r\sin\theta$, z=z.

We presently have at our disposal the following system of differential equations for the light ray trajectory:

(345) $$\frac{d}{dz}\left(\frac{\partial L}{\partial r'}\right) = \frac{\partial L}{\partial r}, \quad \frac{d}{dz}\left(\frac{\partial L}{\partial \theta'}\right) = \frac{\partial L}{\partial \theta}, \quad \frac{d}{ds}(n\tau_z) = \frac{\partial n}{\partial z},$$

of which the first two equations are the ***Euler-Lagrange equations***, and the third is the z component of the ray equation. As we've seen in Chapter 1.2, only two of these equations are independent. We will therefore choose the ones easier to integrate. For the cylindrical distribution considered here, $n = n(r)$, we have $\partial L/\partial\theta = 0$, $\partial n/\partial z = 0$, so from the first two equations at (345) we obtain the conservation of corresponding momenta (impetuses), that is:

(346) $$p_\theta = \frac{\partial L}{\partial \theta'} = \frac{nr^2\theta'}{(1 + r'^2 + r^2\theta'^2)^{1/2}} = nr^2 d\theta/ds = A \quad ,$$

in which:

(347) $$p_z = n\tau_z = ndz/ds = B \quad ,$$

(348) $$A = (p_\theta)_0 = n_0 r_0^2 (d\theta/ds)_0, \quad B = (p_z)_0 = n_0 \cos\gamma_0,$$

are integral constants determined by the initial conditions of the trajectory. In equation (346) we can observe that if $A = 0$, so $d\theta/ds = 0$, then the trajectory remains in a given plane containing the symmetry axis Oz, and constitutes a ***meridional ray***, whereas if $A \neq 0$, so $d\theta/ds \neq 0$, then the trajectory rotates around the Oz axis and constitutes a ***skew ray*** (Fig. 85). We may also observe that any ray launched from the Oz axis ($r_0 = 0$) is meridional, and if $r_0 \neq 0$ the ray is meridional or a skew ray, corresponding to the value of $(d\theta/ds)_0$ being zero or different from zero.

Based on the expression of the path element, equation (342), we also have:

(349) $$\frac{dr}{ds} = \sqrt{1 - \left(r\frac{d\theta}{ds}\right)^2 - \left(\frac{dz}{ds}\right)^2} \quad .$$

We will rewrite equations (346), (347), (349) as the system:

(350) $$\begin{cases} \dfrac{d\theta}{ds} = \dfrac{A}{r^2 n(r)}, \\ \dfrac{dz}{ds} = \dfrac{B}{n(r)}, \\ \dfrac{dr}{ds} = \sqrt{1 - \dfrac{A^2}{r^2 n^2(r)} - \dfrac{B^2}{n^2(r)}}, \end{cases}$$

or, by dividing the first two equations with the third and integrating,

(351) $$\theta = \int \frac{A dr}{r^2 \sqrt{n^2(r) - \left(\dfrac{A^2}{r^2} + B^2\right)}} \quad ,$$

(352) $$z = \int \frac{B dr}{\sqrt{n^2(r) - \left(\dfrac{A^2}{r^2} + B^2\right)}} \quad .$$

Using equations (351) and (352) we've thus obtained the general expression for the trajectory of light rays in media of cylindrical symmetry.

The range of values of r that allow light rays to propagate can be deduced based on the condition that the integral should be a real number, that is, that:

(353) $$n^2(r) \geq \frac{A^2}{r^2} + B^2 \quad ,$$

in which the right term, which through A and B depends on the initial conditions (the equations at (348)), is a constantly decreasing function or r. Let us illustrate by considering the bell-shaped function $n^2(r)$ (Fig. 86). We can see that the light ray can only propagate for the hatched domain $r_{max} \geq r \geq r_{min}$, in which the liminal values r_{min} and r_{max} are the solutions to the equation in condition (353). These limits, which also result from the third equation (350) for $dr/ds = 0$, are the ***turning***, or ***total reflection points***.

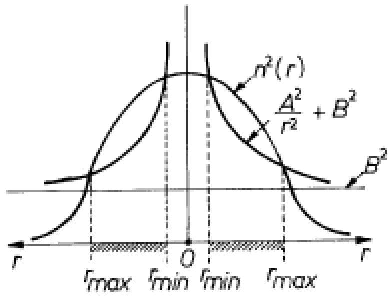


Fig. 86 Bell-shaped distribution $n^2(r)$ and real domains of propagation.

Let us conclude by giving a description of the light guiding property of an ***optical fiber***, featuring a ***Selfoc****-type distribution of the refractive index, namely:

(354) $$n^2 = n_0^2\left(1 - \frac{r^2}{a^2}\right) \quad .$$

By calculating the integrals at (351) and (352) for the distribution at (354), we therefore obtain in principle a skew trajectory in the form of an ellipsis around the symmetry axis Oz (Fig. 87). The projection of this ellipsis across a transverse plane xOy is an ellipsis whose semi-axes are equal to r_{min} and r_{max}, respectively. As the value of constant A increases, the curve $(A/r)^2 + B^2$ in Fig. 86 rises, and the allowed domain $r_{min} \leq r \leq r_{max}$ narrows, until the elliptical helix is reduced to a circular helix ($r_{min} = r_{max}$). If $A = 0$ and $B \neq 0$, the integrand in equation (351) is eliminated, and integral (351) becomes identical to the one that we will obtain in the following Chapter 3.3, describing spherical symmetry, equation (361). In this case, the trajectory will be planar, contained within a plane perpendicular to the symmetry axis Oz.

Note that, in order for the approximation of geometrical optics to be adequate, it is necessary that the diameter of the optical fibers should be larger than the wavelength by a coefficient of at least several score. Because they satisfy this condition, Selfoc optical fibers allow the propagation of light across great distances with minimal losses, and permit important applications in optical communication technology. Similarly, because of their property of periodic light focalization, cylindrical glass bars

* This type of fibers and optical bars are manufactured by the Japanese company Nippon Sheet Glass Co.

with parabolic distribution of the refractive index are used as ***Selfoc micro-lenses***. Such bars of a diameter not exceeding several millimeters (so much larger than Selfoc optical fibers), allow the transferring of images from one end to the other.

3.3 Spherical Structures

Let us first write the equation of the light ray, equation (13):

(355) $$\frac{d}{ds}(n\vec{\tau}) = \nabla n \quad ,$$

multiplied vectorially by the position vector $\vec{r}$. We get:

(356) $$\frac{d}{ds}[\vec{r}\times(n\vec{\tau})] = \vec{r}\times\nabla n \quad ,$$

where we've added the identically zero term $\vec{\tau}\times(n\vec{\tau})$ in order to complete the total derivative of the left member.

In the following lines we will consider a medium whose refractive index depends only on the distance r to a fixed point O, so that $n = n(r)$. For the sake of convenience, we will take the center of symmetry O to also be the origin of the position vector $\vec{r}$ of the light ray trajectory, so that $\nabla n = (\vec{r}/r)\, dn/dr$. Consequently, from the light ray equation written as equation (356) we deduce the fact that vector $\vec{r}\times(n\vec{\tau})$ is conserved along the trajectory, that is:

(357) $$\vec{r}\times(n\vec{\tau}) = \vec{A} \quad ,$$

where $\vec{A}$ is a vector constant for all trajectories. Evidently, this theorem is analogous to that of conservation of kinetic momentum approached in mechanics, regarding the motion of particles within a field of central forces. From equation (357) we deduce that, irrespective of the initial (launch) vectors $\vec{r}_0$, $\vec{\tau}_0$, the light ray trajectory is a planar curve within the plane defined by $(\vec{r}_0, \vec{\tau}_0)$, normal to $\vec{A}$, which also passes through center of symmetry O. Written relative to its absolute value, equation (357) becomes ***Bouguer's theorem***:

(358) $$rn(r)\sin\varphi = r_0 n(r_0)\sin\varphi_0 = r_m n(r_m) = A \quad ,$$

in which φ is the angle between vectors $\vec{r}$ and $\vec{\tau}$, φ_0 is the angle between initial vectors $\vec{r}_0$ and $\vec{\tau}_0$, and r_m is the minimum distance r_{min} (or maximum distance r_{max}) of the trajectory relative to the center of symmetry, so that $\varphi_m = \pi/2$ (Fig. 88). Note that points located at the same distance r from the origin have the same corresponding value $\sin\varphi$, so identical corresponding angles φ and $\pi - \varphi$.

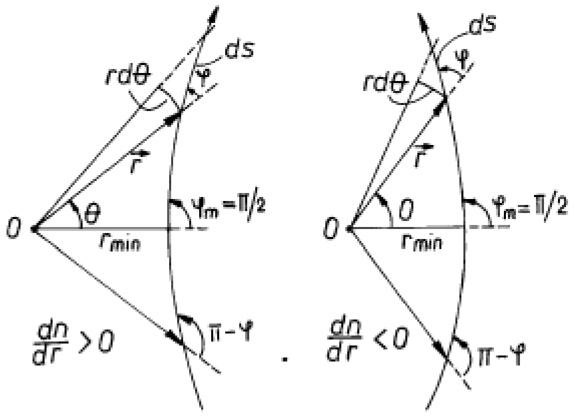


Fig. 88 Light trajectories in structures of spherical symmetry.

Next it will be convenient to discuss the polar coordinates r and θ within the plane containing the trajectory and center of symmetry O. In studying the geometry (Fig. 88), we can deduce the relation $rd\theta/ds = \sin\varphi$, which allows us to rewrite equation (358) as:

(359) $$nr^2\frac{d\theta}{ds} = A \quad ,$$

which is analogous to equation (346) discussed when dealing with cylindrical symmetry. On the other hand, if we use the trajectory path element $(ds)^2 = (dr)^2 + (rd\theta)^2$, we get:

(360) $$\frac{dr}{ds} = \sqrt{1 - r^2\left(\frac{d\theta}{ds}\right)^2} = \sqrt{1 - \frac{A^2}{r^2n^2}} \quad .$$

If we divide the last two equations one by the other and integrate, we obtain the general expression of the light ray path in media of spherical symmetry:

(361) $$\theta = \int \frac{Adr}{r\sqrt{r^2n^2(r) - A^2}} \quad .$$

This is an integral of the same type as the one described previously, in equation (351), where $B = 0$. The domain of values for r that allow the propagation of light rays is obtained by imposing the condition that the integrand should be real, that is:

(362) $$n^2(r) \geq \frac{A^2}{r^2} \quad ,$$

where the right member depends on the initial conditions through constant A,

equation (358), and decreases constantly with r. The ***turning***, or ***total reflection***, points, r_m, are the solutions of the equation in condition (362). Evidently, for these limits we have $dr/ds = 0$, as can be deduced from equation (360).

Let us illustrate by considering first a distribution of the refractive index described by the equation:

$$n^2 = a/r \qquad (363)$$

in which $a(> 0)$ is a constant. In this case the general integral, equation (361), takes the form:

$$\theta = \int \frac{A dr}{r\sqrt{ar - A^2}} \qquad (364)$$

and the trajectory equation is written as:

$$\theta(r) - \alpha = \arccos\left(\frac{2A^2}{ar} - 1\right) \qquad (365)$$

or:

$$\frac{2A^2}{ar} = 1 + \cos(\theta - \alpha) \qquad (366)$$

where α is the integrating constant determined from initial conditions r_0, θ_0. Equation (366) is the polar equation of a parabola of parameter $p = 2r_m = 2A^2/a$ (Fig. 89). So the coordinates of the ***turning***, or ***total reflection***, ***point*** are $r_m = A^2/a$, $\theta_m = \alpha$. Evidently, the distance of the perihelion r_m satisfies (or can be directly calculated by referring to) the general condition, (362).

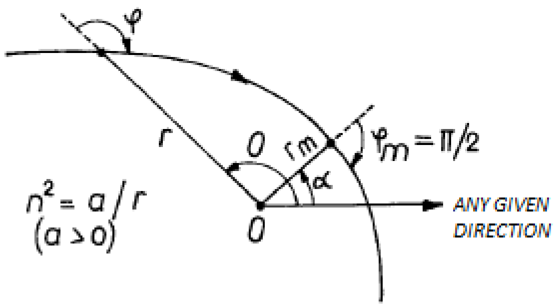


Fig. 89 Turning point (r_m) corresponding to structure $n^2 = a/r$.

A remarkable example of spherical symmetry, first considered by Maxwell (1854), is the distribution termed ***fish-eye***, which is described by the equation:

$$n = \frac{n_0}{1 + (r/a)^2} \qquad (367)$$

in which n_0 and a are constants. In this case, the general integral, equation (361),

becomes:

(368) $$\theta = \int \frac{Cd(\rho - \frac{1}{\rho})}{\sqrt{1 - 4C^2 - C^2(\rho - \frac{1}{\rho})^2}} \quad ,$$

where we've used the following notations:

(369) $$\rho = \frac{r}{a}, \quad C = \frac{A}{an_0} \quad .$$

The trajectory equation thus takes the form:

(370) $$\theta - \alpha = \arcsin\left[\frac{C}{\sqrt{1 - 4C^2}}\left(\rho - \frac{1}{\rho}\right)\right] \quad ,$$

in which α is the integrating constant determined from initial conditions r_0, θ_0. If we invert this relation and go back to the r variable, we obtain the trajectory polar equation written as:

(371) $$\frac{r^2 - a^2}{r\sin(\theta - \alpha)} = 2b \quad ,$$

where constant b is defined as:

(372) $$b = \frac{a\sqrt{1 - 4C^2}}{2C} = \frac{a}{2A}\sqrt{a^2 n_0^2 - 4A^2} \quad ,$$

Next, if we write equation (371) for coordinates $x = r\cos\theta$ and $y = r\sin\theta$, we get:

(373) $$(x + b\sin\alpha)^2 + (y - b\cos\alpha)^2 = a^2 + b^2 \quad ,$$

meaning the trajectories are circles of radius $r = \sqrt{a^2 + b^2}$, with their center at the point of coordinates $x_c = -b\sin\alpha$, $y_c = b\cos\alpha$ (Fig. 90). If $a^2 n_0^2 = 4A^2$, we have $b = 0$; equation (372) and the trajectory constitute the circle of minimum radius $r = a$, centered in the center of symmetry O. Generally, as can be surmised from the polar

equation, equation (371), all trajectories intersect the fixed circle $r = a$ in diametrically opposed points $r = a$, $\theta = \alpha$ and $r = a$, $\theta = \alpha + \pi$ (Fig. 90).

A remarkable property of the fish-eye distribution is that all light rays starting from any given point source $P_1(r_1, \theta_1)$ meet at a point $P_2(r_2, \theta_2)$, so that the coordinates of the two points are linked through the following symmetrical relations:

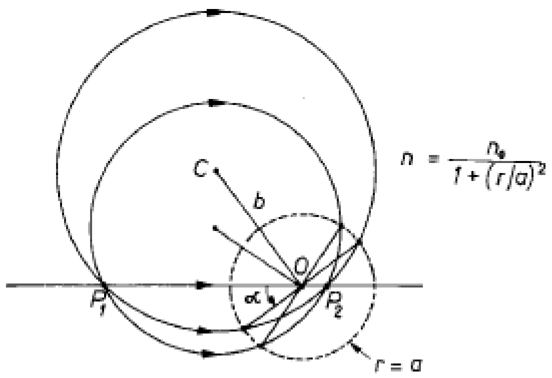


Fig. 90 Maxwell's fish-eye distribution (two circular trajectories).

(374) $$r_1 r_2 = a^2, \qquad \theta_2 = \theta_1 + \pi.$$

Let us demonstrate by describing the group of trajectories (371), of parameter α, passing through point $P_1(r_1, \theta_1)$:

(375) $$\frac{r^2 - a^2}{r \sin(\theta - \alpha)} = \frac{r_1^2 - a^2}{r_1 \sin(\theta_1 - \alpha)} \quad .$$

We can readily verify that no matter the value of parameter α, all the trajectories of this group also pass through point $P_2(r_2, \theta_2)$ of polar coordinates given by equation (374). In other words, points P_1 and P_2, conjugate through the relations at (374), are located on the line passing through the center of symmetry O, one on each side, at distances given by the relation $r_1 r_2 = a^2$ (Fig. 90). Evidently, this geometrical property does not depend on the direction of propagation of the light rays, so that we can just as well consider P_2 as a luminous point source and P_1 as point of focalization. The fish-eye distribution is a classic example of a ***perfect optical instrument***, in the sense that any conical (homocentric) light beam projecting from a random object point in space is transformed into a conical beam that converges in the corresponding image point.

Another interesting example of spherical distribution is the ***Luneburg lens***, comprising a non-homogeneous sphere of unit (relative) radius and refractive index $n = \sqrt{2 - r^2}$ (for $r \leq 1$), placed in air ($n = 1$ for $r > 1$): Fig. 91. Because for $r = 1$ we have $n = 1$ (on the outside, as well as on the inside), no ray of incidence will be refracted per se at the surface of the lens; refraction takes place continuously within the lens (for $r < 1$).

In the case of this type of lens, relation (361) yields us:

(376) $$d\theta = \frac{A\,dr}{r\sqrt{r^2(2-r^2)-A^2}} \quad .$$

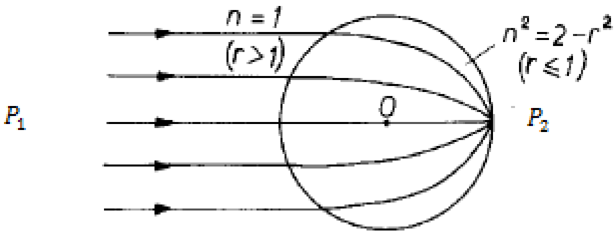


Fig. 91 Luneburg lens.

If we use the notation i for the angle of incidence to the lens (where $r = 1$ and $n = 1$) of a light ray originating in P_1 (located at infinity) and refer to relation (358), we discover that $A = \sin i$, and so:

(377) $$d\theta = \frac{\sin i\,dr}{r\sqrt{\cos^2 i-(r^2-1)^2}} \quad ,$$

After integrating and considering the initial condition $\theta = i, r = 1$, we obtain the trajectory:

(378) $$\theta - i = \frac{1}{2}\arcsin\left(\frac{r^2-\sin^2 i}{r^2\cos i}\right) - \frac{1}{2}\arcsin(\cos i) \quad ,$$

or

(379) $$\frac{1}{r^2} = \frac{1}{\sin^2 i}[1-\cos i\cdot\sin(2\theta+\arcsin(\cos i))] \quad .$$

From here we may surmise that, for ***any*** i, at $\theta = \pi$ we have $r = 1$, meaning ***perfect focalization*** in point P_2 on the sphere surface, as can be seen in Fig. 91. Consequently, the image of an extended object located at a long distance from the lens will form across a spherical surface of unit radius. This kind of lenses manufactured from porous plastic material have been used in the field of microwaves. They differ from the regular lenses of optics, in whose case refraction only takes place on the surface.

A more general continuous distribution of the refractive index that allows the achievement of stigmatism is described as:

(380) $$n^2r^2 = n^{1/p}(2-n^{1/p}), \qquad (r \le 1),$$

where $p > 0$. Particularly, for $p = 1/2$, we obtain the distribution of the Luneburg lens. This kind of distribution is useful in the field of short electromagnetic waves (decimetric and centimetric), in constructing projecting systems that transform a homocentric beam originating in a point source into a parallel beam (planar waves), but also in the field of optics where through its two-dimensional versions it is used as

lenses in integrated optical circuits. Fig. 92 illustrates a ***two-dimensional Luneburg lens*** (seen from above), which transforms concentric circular objects into images and vice-versa.

In the examples shown above we've considered the problem of determining the trajectory $\theta(r)$ of light rays, equation (361), for a given distribution $n(r)$. Naturally, the problem can be formulated the other way around: determine the distribution $n(r)$ from the trajectory integral equation, relation (361), considering given light ray trajectories $\theta(r)$. For example, let us determine the nature of the dependency $n(r)$ of the refractive index of a spherically symmetrical medium, so that the light ray trajectory $r(\theta)$ should be a parabola (conical, of eccentricity $\varepsilon = 1$) of equation:

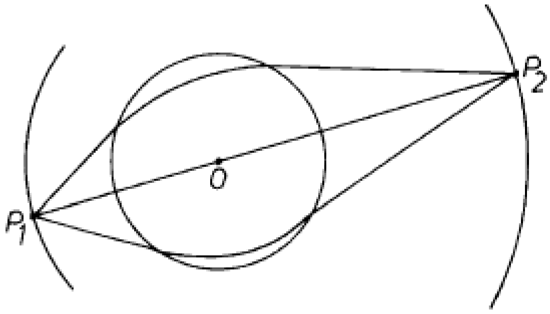


Fig. 92 Luneburg two-dimensional lens and conjugate circles.

(381) $$r = \frac{p}{1+\cos\theta} = \frac{p}{2\cos^2(\theta/2)} \quad , p = \text{known.}$$

Referring to equation (361) we immediately obtain:

(382) $$n(r) = \frac{A}{r}\left[1 + \left(\frac{1}{r}\frac{dr}{d\theta}\right)^2\right]^{1/2} \quad ,$$

in which we then make the following replacement:

(383) $$\frac{1}{r}\frac{dr}{d\theta} = \sqrt{\frac{2r}{p} - 1} \quad .$$

We thus obtain:

(384) $$n(r) = \frac{A\sqrt{2}}{\sqrt{pr}} \quad ,$$

where, according to relation (358), $A = r_{min} n(r_{min})$, in which $r_{min} = p/2$. If we use the notation $a \equiv 2A^2/p$, the result we thus reach coincide with dependency (363) previously discussed.

Readers interested in the issues discussed in this chapter will find additional information in works [16], [29], [45], [46], [54], and [95].



Appendix A
MOMENTS IN THE HISTORY OF GEOMETRICAL OPTICS

Archeological research reveals that, as far back as 4000 years ago,
around 2000 BCE the ***Egyptians*** mastered the ***technique* of *polishing metallic mirrors***
made from copper, bronze and later speculum, a copper alloy, rich in
tin. One such mirror, dated around 1900 BCE, was discovered in perfect
around 1000 BCE condition near the pyramid of Sesostris II along the Nile river valley.
Archeological evidence also proves that ***rudimentary lenses*** were
already in use between 3000 and 3500 years ago. First written mention
of a ***converging lens***, used to focus sunbeams, may be found in
424 BCE ***Aristophanes***' comedy ***The Clouds*** of 424 BCE, in which a certain debtor
is thus able to destroy the proof of his debt, marked on a wax tablet, from
afar.

The great Greek philosophers speculated much on the nature of
light and the mechanism of sight, and advanced simple hypotheses that
today seem partly strange and partly essentially correct. The geometer
500 BCE ***Pythagoras*** (582 – 500 BCE) believed that the human eye emitted light
beams, like a headlight, by use of which it "felt" surrounding bodies, a
principle foreshadowing today's radar and sonar devices (!); ***Empedocles***
(490 – 430 BCE), who authored the doctrine of matter being composed
of particles of the four elements (earth, water, air, and fire; today we
would say solids, liquids, gases, and plasma), advanced the hypothesis
400 BCE that light propagates through space at a finite speed; ***Democritus*** (460 –
370 BCE), the father of the atomist doctrine, supposed that visual
sensations are caused by minute material particles (named eudols)
emitted by objects. ***Plato*** (427 – 347 BCE), author of the famous dialogues
and advocate of the "ocular rays," makes the first important observation
regarding the refraction of light (in the Republic, line 602, book X), and
Aristotle (384 – 322 BCE), the greatest scientific figure of Antiquity,
objects to light beams being emitted by the eye and advances an ether
hypothesis similar to that of the 19th century.

300 BCE The great Greek geometer ***Euclid*** of Alexandria, from around 300
BCE, author of Antiquity's mathematical work ***Elements (of Geometry)***
in 13 volumes (see ***Elements of Geometry***, three volumes, Dover
Publications Inc., New York, 1956), also wrote the first great work of

optics, the ***Optics***, in which he sets the foundation for perspective theory. Also sometimes attributed to Euclid is a work on ***Catoptrics***, which contains a study of the laws of reflection, a series of theorems on planar mirrors, and a description of the focalizing action of concave mirrors. It follows that Euclid quite probably was aware of the laws that up until this day mark the foundation of Catoptrics, which state that: (1) in homogenous media light propagates in a straight line, (2) the angles of incidence and reflection are equal, and (3) the rays of incidence and reflection belong to a plane perpendicular to the mirror's surface. However, it must be said that, unlike Euclid's work on geometry, which exhibits flawless logic, his work on geometrical optics also contains many inaccuracies. It is doubtful whether these assertions were authored by the great geometer.

212 BCE ***Archimedes*** of Syracuse (287 – 212 BCE), Greek mathematician and physicist, considered to have been the greatest mathematical genius of Antiquity, founder of statics and hydrostatics, is mentioned in Roman historiography in the legend surrounding his using mirrors spread along the shore to reflect solar radiation onto the Roman fleet that besieged Syracuse (212 BCE).

50 BCE ***Cleomedes*** (around 50 BCE), Greek astronomer, describes the refraction of light by demonstrating that a light ray approaches the normal line when it enters a denser medium and moves further away from it when it enters a medium of lower density. He affirms that the Sun is visible even below the horizon, thanks to atmospheric refraction. He mentions the "coin in a cup" experiment made by Ctesibius at the University of Alexandria around 50 BCE, which consisted in placing a coin on the bottom of an empty cup and making the coin visible to surrounding observers by filling the cup with water.

Seneca, Lucius Annaeus (around 4 BCE – 65 BCE), Roman philosopher and politician, observed that a globe made of glass could be used to enlarge images. It is quite possible that Roman craftsmen even then used enlarging lens for very fine work. Also, ***Caius Plinius Secundus***, or ***Pliny the Elder*** (23 – 79) was aware that a glass sphere could kindle certain matter placed in its focal point when exposed to the Sun, and draws attention to his application of this device to cauterizing wounds. The well-known roman naturalist (author of an encyclopedia on natural history comprising 37 volumes) died while closely observing



an eruption of Mount Vesuvius (79 CE).

50 CE ***Hero*** (***Heron***) of Alexandria (probably 1st century), Greek mathematician and inventor, great experimenter, is known in particular for his water, vapor or compressed air based machines and devices. Heron's ***Catoptrics*** is one of the most interesting books on optics in Antiquity. In this work, Heron observes a deeper reason underlying the laws of catoptrics (see the paragraph on Euclid above), by postulating that ***light rays propagate from one point to another along the shortest possible route***. This assertion constitutes the first formulation of a natural law using a variational principle. The same idea was later taken up by Fermat (1657), who generalized it under the form of the "principle of least time" to also encompass an explanation for the law of refraction. Within the same work, being the remarkably imaginative experimenter that he was, Heron also described numerous amusing or practical effects one could achieve by the use of planar and cylindrical mirrors.

150 ***Ptolemy, Claudius, of Alexandria*** (100 – 160), Greek astronomer, mathematician and geographer, became famous for his work "Mathematical Treatise" (Almagest), and for his Ptolemaic geometrical system – a geometrical design devised by Ptolemy by which the apparent movements of the Sun, Moon and planets, as seen by a terrestrial observer, are predicted with a high degree of precision (with a deviation of two degrees at most from the observer position). According to this design, the movement of the heavenly bodies result from the composition of uniform circular motions, namely a circular motion based on a small circle (the epicycle), whose center moves along a large excentric circle (the deferent) around the Earth (see S. Olariu, ***Geneza și evoluția reprezentărilor mecanicii clasice,*** The Scientific and Encyclopedic Press, Bucharest, 1987). Ptolemy also possessed the most extensive and thorough knowledge of optics in Antiquity, and wrote one of the most remarkable treatises on optics, ***Ptolemy's Optics***. For a long time believed to be lost and known solely through citations found within writings from the Middle Ages, it was ultimately recovered from the latin manuscripts, ***Ptolemaei opticorum sermones quinque***, translated from Arab. This treatise encompasses all the branches of optics that were then known, namely sight, reflection on planar and concave mirrors and refraction. Ptolemy also undertook a serious experimental study of the refraction of light, a singular work in Antiquity, in which he studied the

refraction of light through air into water and glass, and through glass into water, he described the measuring instrument (a circular graded disc for measuring the two angles), and left tables containing the experimental data regarding angles of refraction corresponding to angles of incidence (every 10 arc degrees). The data regarding the 50° and 60° angles of incidence is surprisingly of particular precision. It is evident that Ptolemy could have discovered the exact law of refraction. Unfortunately, it remained undiscovered until Snell (1621) and Descartes (1672). Ptolemy also does not mention anything regarding the separation of color through refraction (light dispersion), leaving this phenomenon to be discussed by Newton (1672). On the other hand, his astronomical preoccupations led to his learning more about astronomical refraction. Thus, Ptolemy discovered that the apparent position and actual position of a star coincide only when located at the zenith, and that the rest of the sky showed the heavenly bodies at a greater height than their actual position due to atmospheric refraction, the discrepancy becoming greater the closer the heavenly bodies are to the horizon. In other words, he knew that light rays entering the atmosphere obliquely curb towards Earth.

This is the extent of the knowledge the scholars of Antiquity had in the field of optics. Although they are not impressive, they far surpass the knowledge held at that time in any other branch of physics.

Abu Ali Al-Hasen ibn Al-Hasan ibn Al-Haytam, or ***Alhazen***, for
1025 short (around 965 – 1039), Arab mathematician and physicist, the greatest scholar of optics of the early Middle Ages, author of the treatise on optics entitled ***Kitab Al Manazir***, made valuable contributions to geometrical and physiological optics. Alhazen extended the research regarding the reflection of light to encompass conical, concave, and convex surfaces, he formulated the problem of the position of the point of incidence of a light ray for the given positions of the eye and the luminous point (Alhazen's problem), and he postulated the law of reflection, observing that the ray of incidence, the reflected ray, and the normal line to the mirror surface in the point of incidence are located in the same plane. He also extended this observation to the law of refraction (his experiments are still used today for illustrative purposes), he remarked that Ptolemy's observation, that the angles of incidence and refraction are proportional, is only valid for sufficiently small angles, but



he, too, missed the opportunity to discover the exact mathematical formula for this law, although the Arab mathematicians had already developed the concept of the sine. Alhazen conclusively disproved the "ocular ray" hypothesis inherited from Pythagoras and Plato and still present in Ptolemy's ***Optics*** using factually based arguments, namely that sight is drastically influenced by external conditions such as lighting and the apparent color, dimensions, and intensity of light given off by objects. The only simple explanation would then be that sight is caused by something that travels from objects to the eye. By observing that the effect of intense light is felt even after closing one's eyes, Alhazen concluded that light induced certain reactions within the eye. In order to understand the mechanism of sight, he conducted a detailed anatomical study of the human eye, correctly describing the cornea, choroid, iris, crystalline lens, humors and retina, with its nervous structure. This became a classical description. Even the term *lens* comes from the Latin translation ***lens*** of what Alhazen in Arabic called ***adasa***, meaning ***bead***, when he described the crystalline lens. He later constructed a simple physical model of the eye, the famous pinhole camera, by the use of which he conducted numerous experiments. However, the fact that the image was inverted in the pinhole camera led him to believe that, by comparison with the eye, the image was perceived on the first surface of the lens, and not on the retina. There would still be a long while before Kepler (1600) would realize that the retina is the photosensitive stratum after all, and that setting the inverted image right was a physiological effect. Alhazen also used the pinhole camera as a veritable precursor of the photographic camera in order to study solar eclipses. The work of the great Arab scholar was translated into Latin as ***Opticae Thesaurus Alhazeni***, and had a profound influence on Robert Grosseteste, Roger Bacon, Witelo, Leonardo da Vinci, Johannes Kepler and Isaac Newton.

1200 ***Grosseteste, Robert*** (1168 – 1253), British philosopher, the first rector at Oxford University, was one of the European pioneers of the experimental and deductive method. He wrote about movement, heat, sound and light, he carried out experiments using mirrors and lenses, and considered optics to be the foundation of all sciences.

1270 ***Bacon, Roger*** (1214 – 1294), British philosopher, compiled all his writings on optics in ***Opus Majus*** (1267). Among his achievements, the following are worth mention: he precisely determined the focal point of



spherical concave mirrors, he described the spherical aberration, and recommended a parabolic design for mirrors. Roger Bacon was a man of great erudition, and was known among his contemporaries as ***Doctor mirabilis***. He was endowed with an original spirit and the inventive genius, which would have perhaps led him to important discoveries if he had lived in more enlightened times and more favorable circumstances. It is known that his work abounds in projects which he never finished (road and sea transportation, flight, submarine exploration, medical remedies obtained using chemicals, even gun powder, which was invented much earlier!). In particular, within the field of optics, he intuitively guessed the possibility of constructing the magnifying glass and the telescope, and he measured the angle of the rainbow. However, his writings are so vague and brief that it is impossible to attribute to him the authorship of any one invention. Despite all this, Roger Bacon is considered to have been an early proponent of modern science due to his very wide-ranging interest for science and to his open-minded and firm belief that useful knowledge cannot be obtained through baseless speculation, but only through discovering the facts based on the solid foundation of observation, experiment, and mathematical reasoning.

Witelo, lat. Vitellius (1220 – ?), Polish physicist and philosopher,
1274 studied in Italy (1269). Through his work on optics (from around 1274), which was voluminous, but contained little new knowledge, he established a connection with the Greek and Arab optics, in that he again brought to attention what Alhazen and the scholars of Antiquity had left in this field. Unfortunately, Witelo ultimately became famous for his errors related to his experimental calculation of the angles of refraction of light passing through air into glass and water, which yielded impossible results (they were at odds with the phenomenon of total reflection, which was not yet known at the time). It is possible that this very data later led Kepler (1611) astray, preventing him from finally discovering the exact law of refraction. We would be unjust not to acknowledge Witelo's merits, beginning with his volume on optics, which he gave to Europe. Empirically, he was aware of the phenomenon of dispersion, which always co-occurs with the refraction of light, he made the natural observation that, in refraction, the two angles remain the same whichever way light propagates, he noticed that some of the

light is lost in reflection and refraction, he contributed to the psychology of sight, and came very close to understanding the phenomenon of rainbows.

1285 ***Degli Armati, Salvino*** (? – 1317), Florentine noble, is apparently the inventor of the ***eyeglasses***, according to an epitaph in Florence. In reality, the identity of the inventor of the glasses is a problem shrouded in mystery, and whoever it was, he/she probably had nothing to do with the underlying theoretical aspects, but more likely made this achievement through practice and on account of chance. Locating the moment of invention within the 13th century is supported by the dictionary of the Accademia della Crusca, in which the year given for the invention is 1285, and by an old manuscript dated 1299. The first portrait of a man wearing glasses was painted by ***Tomasso di Medina*** (1352). Only convergent lens for long-sightedness were initially used. There is no reference to the use of divergent lens for short-sightedness before the second half of the 15th century.

1485 ***Leonardo da Vinci*** (1452 – 1519), Italian artist and scientist, considered to be the greatest observer of nature of all time, whose passion for art, painting, sculpture, architecture, and music led him to vast scientific and technical research. He observed the resistance, compressibility, and weight of air, as well as the flight of birds, and designed the glider and the parachute. He studied the shapes formed by sand on vibrating surfaces, standing waves in liquids, friction, the effective weight of bodies placed on inclined planes and the parallelogram of forces, he designed canals, irrigation systems and artistic constructions, he invented the elevator, and researched the distribution of tension in arcades, columns, and walls. He studied capillarity and the formation of drops, and compared the density of liquids by balancing them in a U-shaped tube. The quality of his research and observations concerning human anatomy were unmatched. He studied the anatomy of the eye and devised a model for it (he assumed that light rays traverse it twice in order for the image to be upright!), he observed the phenomenon of diffraction, and he designed a photometer and a device for polishing concave mirrors. Unfortunately, with the exception of his ***Trattato della pittura***, most notes and designs belonging to this man of genius were left in unorganized manuscripts, of which some were lost and some have remained unknown until

recently (Venturi, ***Essai sur les ouvrages de L. da Vinci***, 1797).

1575 ***Maurolico, Francesco*** (1494 – 1575), Italian geometer and scholar in optics, of Arab descent, known for his well-written work on optics, entitled ***Photismi (theoremata) de lumine et umbre*** (Venice, 1575). He surpassed his predecessors, Alhazen and Witelo, in that he no longer located perception of images in or on the crystalline lens, but behind it, likening its functioning to a biconvex lens, thus explaining short and long-sightedness. In order to give an idea of why understanding the workings of eye was so difficult, it would be useful to mention that Maurolicus had not yet realized that a real image forms on the retina in the back of the eye. He also observed, for the first time, that light rays given off by a punctiform source have a wrapping surface known today as ***caustic***, on which the concentration of light is at its maximum level, and he measured the angular diameters of the rainbow arcs.

Porta, Giambattista della (1534 – 1615), Italian naturalist, wrote one of the most ″colorful″ works of his time, ***Magiae naturalis libri XX,***
1589 ***1589***, later translated into five languages (Italian, French, Spanish, German, and Arab), a veritable vade mecum of the times, a bizarre mixture of useful recipes and myths and legends gathered uncritically from the most diverse sources, which the author refrains from specifying: from cosmetics, perfumes and distillation, gardening, housekeeping and obtaining wealth, pyrotechnics, metallurgy and fake precious stones, to astrology, sympathy-antipathy, and palmistry. It is in only in his seventeenth volume that he discusses optics, in which he presents numerous tricks using mirrors, as Hero had done 1500 years before, but also the first exact theory of multiple mirrors, the first complete description of the camera obscura, with pinhole or with lens, he compares the eye and pupil to the camera obscura and diaphragm, and he presents various combinations of convergent and divergent lenses, which later led to his claims to having invented the telescope.

For indeed, the 1590 – 1610 moment of the telescope and microscope had finally arrived. These inventions were achieved empirically by the Dutch opticians ***Hans Lippershey***, ***Jacob Adriaanszoon***, or ***Hans*** and ***Zacharias Jansen***, and out of scientific interest by ***Galileo Galilei*** and ***Johannes Kepler***. But let us first see the state of astronomical observation in the period leading up to the discovery of the refraction telescope.

Brahe, Tycho (1546 – 1601), Danish astronomer, perfected the art of astronomical observations with the naked eye, thus bringing the art to its limit, to this end using large and precise mechanical devices. For example, Tycho Brahe's quadrant had a radius of almost two meters and a greater level of precision than five arc seconds. In order to get an idea of the magnitude of this achievement, we note that a man in Bucharest is seen in a straight line from Drobeta-Turnu Severin under an angle of one arc second. Tycho Brahe's data regarding the positions and movements of the stars, Sun, planets, and Moon, whose precision astonishes us even today, was published in ***Astronomiae instauratae***
1598 ***mechanica*** (1598). It allowed Kepler (***Astronomia nova***, 1609) to discover that the planets' orbits are ellipses, with the Sun being located in one of the focal points, and stimulated Römer's studies, which led to his determining the speed of light (1676).

In Holland, as in Italy (near Venice), quality glass was produced for eyeglass lens. Although Holland is undeniably the country in which the ***telescope*** was invented, it is not at all clear who exactly the inventor was. According to the Hague archives, the first request for a patent for
1608 this instrument is dated 20th of October 1608, and was made by the optician and lens craftsman **Hans Lippershey** (1587 – 1619), but was immediately contested by his fellow craftsmen ***Jacob Adriaanszoon***, and ***Hans*** and ***Zacharias Jansen*** (his son). In the resulting confusion, the right of patent requested by Lippershey was denied. It is known for a fact that Dutch spyglasses were being sold in 1609 in Paris, Frankfurt, London, Milan, and Padua. It is not surprising that the ***compound microscope*** was invented practically at the same time, being attributed to ***Zacharias Jansen*** (1588 – 1632) and his father, although several others claimed precedence in this case as well (George Huefnagel of Frankfurt and the astronomer Francesco Fontana of Napoli). Thus, around the year
1610 1610, the notion that using two lenses would produce an image much larger than that obtainable using one was put in practice. Use of the new optical instrument for seeing across large distances – the refraction telescope – spread rapidly throughout Europe. It is understandable why the authorities had a special interest in this device. Its impact on science was immediate, as Galileo and Kepler would soon demonstrate.

Galilei, Galileo (1564 – 1642), Italian astronomer, mathematician and physicist, professor at the universities of Pisa and Padua, for his



fundamental discoveries in mechanics and astronomy, he is considered to be the founder of exact sciences and the modern scientific method. His descriptions of his experiments and mathematical reasoning written following this principle are comparable to those written today, as we can observe by reading his useful work ***Discorsi e dimostrazioni matematiche intorno a due nuove scienze*** (Leiden, 1638), translated into Romanian by Victor Marian, with the title ***Dialoguri asupra științelor noi***, Academy Press, 1961. Firstly, we note that the way in which Galileo studied uniformly accelerated movement represented a veritable starting point for differential representations later introduced in physics by Newton (1669). Among his achievements in the field of mechanics we mention the discovery of the isochronism of pendulum oscillation (1583), the invention of the hydrostatic balance (1586), his classical experiments and theory regarding falling weights and the movement of projectiles (see ***Dialoguri***, mentioned above). Yet Galileo's fame among his contemporaries was due to his crafting of the first ***practical telescopes*** (1609) and the astonishing astronomical discoveries he made using such a device, beginning with the nights from the 7th to the 15th of January 1610, when mankind finally opened its eyes to the heavens. In fact, Galileo first heard about Dutch telescopes as late as May 1609, while he was professor of mathematics in Padua, and had not yet begun his study of optics. These telescopes, consisting of a convergent objective lens and a divergent eyepiece lens, were very rudimentary and did not exceed a magnifying power of 3X. Galileo quickly set to single-handedly fashioning his own telescopes, by fashioning two lenses, one convex and one concave, which he fixed within an organ pipe, an instrument which he improved in time, quickly reaching magnification powers of 14X, 20X, and finally 30X (he crafted over 100 within his lifetime). In August of 1609, Galileo was already presenting the Senate of Venice with such a telescope, much more powerful than the one he had received from Holland, and in January of 1610 directed his telescopes towards the sky, finding heavenly bodies never before seen. He thus observed the four main satellites of Jupiter (three on the night of January the 7th, the fourth becoming visible on January the 13th), a veritable miniaturized Copernican system (the orbits of these satellites are located within the plane of Jupiter's equator, their semi-axes having a length equal to several of the planet's diameter, the satellites' orbital periods having a

duration of several days), he discovered the lunar mountains, valleys, and craters (the surface of the Moon was therefore not smooth, as the ancient Greeks had supposed), the ring surrounding the planet Saturn (initially named ***planeta tricorporeum***), the phases of Venus, sunspots and the rotation of the Sun, and the Milky Way's star structure (in which he saw thousands of stars, invisible to the naked eye). The first of these
1610 observations he published in ***Nuncius sidereus*** (Venice, March 1610), sparking incredible excitement. The scientific impact of these discoveries, made by Galileo with the telescopes he had crafted, which revolutionized man's view of the Universe, was immediate, as can be easily ascertained from the enthusiasm with which a man of Christian Huygens's repute writes about them in ***Dioptrica, de telescopiis***, 1653. From the invention of the telescope (objective of large focal distance), the invention of the microscope (objective and eyepiece of small focal distance) was only one step away, which Galileo also took (1610), along with the Dutch inventors, thus becoming one of the precursors of the modern compound microscope. The hardships that Galileo went through for his discoveries, on account of ignorance and the Inquisition, are well known, but the burning flame of reason and science could no longer be quenched, because it was already the onset of the 17th century, the century of scientific giants: Galileo, Kepler, Descartes, Pascal, Fermat, Newton, Leibniz, and Huygens.

Kepler, Johannes (1571 – 1630), German astronomer, assistant (1600 – 1601) and then successor to Tycho Brahe at the Prague observatory. Driven by a passion for precision, and using Tycho Brahe's exact observations, as well as employing great care and perseverance in his calculations, which he redid countless times, Kepler uncovered the laws governing planetary orbits, which he formulated in his work
1609 ***Astronomia nova***, Prague, 1609, thus putting an end to a 2000-year-old tradition of circular geometrical representation (see S. Olariu, ***Geneza și evoluția reprezentărilor mecanicii clasice***, The Scientific and Encyclopedic Press, Bucharest, 1987).

Kepler did not have the same kind of luck when he used Witelo's erroneous experimental data regarding angles of refraction, from which it may be deduced how important the precision of observations and measurements is to gaining correct knowledge. In his first work on optics, entitled ***Ad Vitellionem paralipomena, quibus astronomiae pars***

1604 ***optica traditur***, Frankfurt, 1604, Kepler thus attempts, and almost succeeds in discovering the exact law of refraction. To this end, he began by considering the issue of Cartesian surfaces, 33 years before Descartes wrote ***La Dioptrique*** in 1637, for the particular case of a surface that would refract a parallel beam of light so that it converges into a punctiform image. He had the inspiration to choose a hyperboloid surface, which, as we know, meets the requirements of the exact law of refraction, but, unfortunately, since he trusted the data found in Witelo's tables, he gave up the project, because his surface did not correspond to said erroneous data, which expressed a false relationship between the angles of incidence and refraction.

Galileo was in correspondence with Kepler, whom he also presented with a telescope he had fashioned (1610), by use of which he had recently inaugurated the era of great astronomical discoveries. Stimulated by these achievements, Kepler resumed his optical research, and wrote his second book on the subject, entitled ***Dioptrice***, Augsburg,
1611 1611, with which he gave significant impetus to the progress of a science whose theoretical aspects had been too long neglected. In this written work, which comprises 80 pages, and which may be fruitfully read even today, based on simple considerations of elementary geometry, Kepler laid the ***fundamental principles of dioptrics*** within the frame of the paraxial approximation (in which the angles of incidence and refraction are proportional), with application to thin lenses and doublets, as well as triplet lenses. In this context, Kepler formulates ***the first theory of refraction telescopes*** (for which the paraxial approximation is excellently appropriate, since, in practical applications, the light rays' angle of incidence is always very small), and describes five types of telescopes, namely: (1) the convex objective/concave ocular doublet (the Dutch telescope, or ***Galilean telescope***), (2) the convex objective/convex ocular doublet (a system advanced in sentence 88 of the ***Dioptrice***, which bears the name ***Keplerian telescope***), (3) the convex objective/double concave ocular triplet, (4) the convex objective/double convex ocular triplet, and (5) the double convex objective/convex ocular triplet. The second telescope described, which Kepler suggested in his work, would
1613 be in fact constructed several years later (around 1613) by his compatriot, the astronomer ***Cristoph Scheiner*** (1575 – 1650), who for several years would use the invention in order to make the first systematic

observations on the movements of sunspots (results which he published in his work ***Rosa Ursina***, 1626 – 1630). Unlike Galileo's telescope, Kepler's telescope was used for astronomical observations (to which the fact that the image appeared inverted did not constitute an impediment) due to its relatively large angular field of view. Furthermore, Kepler's telescope had the essential advantage of forming a real intermediate image, so that, by attaching across its plane two reticular threads intersecting along the telescope's axis, or a graded circle, the telescope could be used for precise positioning and measuring. From among the people who played an important role in the history of the telescope's invention, we would also like to mention ***Fontana*** and ***Schyrl*** (or ***Schyrleus***). Italian astronomer ***Francesco Fontana*** (1580 – 1656), the author of ***Novae celestium terrestriumque rerum observationes***, 1646, first person to observe the "channels" of Mars, and probably first to change the initially concave microscope ocular with the convex ocular, also claimed to have had invented the astronomical telescope (Keplerian telescope) as far back as 1608. More certainly recognized are the merits of astronomer ***Anton Maria Schyrleus of Rheita*** (1597 – 1660), the
1645 inventor of the ***terrestrial telescope*** (1645), a convex lens quadruplet telescope, or double astronomical telescope, which renders upright images of objects. He is also credited with coining the terms ***objective*** and ***ocular***. ***Demisiani***, a member of the Accademia dei Lincei, is the one who coined the names ***telescope*** and ***microscope***, which replaced the words ***conspicilia***, ***perspicilia***, ***occhiali***, and ***occhialini*** that had been used up to that moment.

Besides introducing the theory of dioptrics and telescopes, Kepler also determined the correct path light rays travel within the eye, by demonstrating that each cone of rays emitted by the points on the object is refracted through the crystalline and recombined into the corresponding image point on the retina; sight is thus the sensation of retinal stimulation, and thus the analogy between the eye and the camera obscura with its convergent lens is correct. All this ***theory of sight*** was formulated by ***Kepler*** in his first work on optics of 1604, and was soon afterwards confirmed experimentally by Scheiner (who constructed the above mentioned Keplerian telescope) in a remarkable
1610 work, ***Oculus, hoc est fundamentum opticum***, 1610. Scheiner wrote that the refractive index of the aqueous humor is equal to that of water, and

that the index of the vitreous humor is approximately equal to that of glass; it is also in this work that he describes his famous cow's eye experiment. By removing the posterior sclera and choroid, and looking through the back of the eye as if through the objective of a telescope, he very clearly saw the image of distant objects on the transparent retina. Later on, in 1625, he made the same demonstration using a human eye. This decisive and quite enlightening experiment proved once and for all that the ***locus of sight is on the retina***. Scheiner simultaneously explained ***how the eye adapts*** by modifying the shape of the crystalline lens (an effect which he obtained in the abovementioned experiment by gently pressing on the eye), which becomes more convex (more curved) when looking at closer objects, and less convex (more flattened) when looking at objects farther away. We are approaching a culminating point in geometrical optics, when ***the simple and exact law of refraction***,

$$\frac{\sin i}{\sin r} = n \,,$$

was once again "in the air". It had once floated above the Greek astronomer and mathematician Ptolemy (150), had circled around the Arab mathematician and physicist Alhazen (1025), and had slipped through the hands of Polish physicist Witelo (1274), as we've shown in the previous pages. Now it was freely presenting itself to the great discoverer of laws himself, Kepler, in the form of a pressing need to finally understand the functioning mechanism of these strained achievements of the empirical, the eyeglasses, microscope, and telescope. All that was needed for it to be discovered was the sufficiently precise measuring of ***two angles***: the angle of incidence i, and the angle of refraction $\boldsymbol{r}$. And these measurements were once again taken by
1611 Kepler (***Dioptrice***, 1611), who, however, only ended up with the same results inherited from Ptolemy, that for angles of incidence not exceeding around 30°, it can roughly be said that $i/r = n$, where, for light traversing through air into glass, n = 3/2. However, he also discovered ***the critical angle of total reflection***, which for glass-air reflection is around 42°. From here onward, Witelo's erroneous data conspicuously contradicted experience. Unfortunately, Kepler never recalled the issue he had approached seven years earlier, in his first work

on optics of 1604. And so, Kepler, too, was not fortunate enough to discover the true law of refraction, in spite of the opportunity which had presented itself, a surprising development for a scholar of his stature.

But time had lost its patience, and the capricious law wrote itself (and, essentially, using no other means than those that had been available for almost 1500 years) through the hands of ***Snell*** (around
1621 1621), in the form ***cosec r/cosec i = n***, and through Descartes (1637), in
1637 the form ***sin i/sin r = n***. Thus, fig. A.1, considered within the plane of the ray of incidence AI and the ray of refraction IB, in the case of air-glass propagation (n=1.5), illustrates both "expressions" of this law, namely Snell's formula, who wrote ***IB/IE = n***, and Descartes's version, who wrote ***IC/ID = n***.

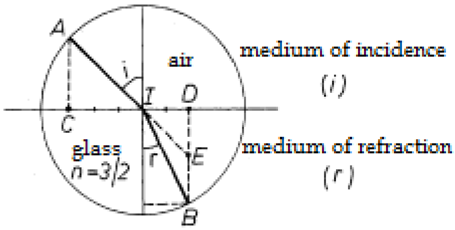


Fig. A.1 The law of refraction

Snell, Willebrord van Roijen, lat. ***Snellius*** (1591 – 1626), Dutch mathematician, professor at Leiden, known for his works on spherical trigonometry and triangulation, also wrote a work on optics, in which he set forth the law of refraction in the form of the ratio between cosecants, its value in the case of air-water refraction being correctly calculated as ***IB/IE = n = 4/3***. It is plausible to assume that Snell obtained this law using more precise angle measurements, and that he drew the idea of expressing it in this form from ***Ctesibius's*** "coin in a cup" experiment (the apparent rising of the bottom of a container filled with water, 50 BCE). The fact is that no one knows how Snell thought of, and determined, the law that opened the way for modern optics in one fell swoop, because his writings on optics were never published. It would be interesting to imagine what our scientific and technological evolution might have been if this path had been set by Ptolemy, Alhazen, or at least Witelo. As Huygens has remarked, it seems that Snell did not really understand the importance of his discovery. Fortunately, Huygens, who read Snell's manuscript (as did Isaac Voss, to whom we will refer presently), fittingly cited it in his first book on optics, ***Dioptrica***, of 1653.

Descartes, René du Perron, lat. Cartesius, Renatus (1596 – 1650), French mathematician, the founder of geometric analytics, wrote an interesting book on optics, ***La Dioptrique***, 1637, in which he expressed

the exact law of refraction in its actual form of sinus ratio. The way in which Descartes tried to explain this law using an ad hoc mechanical analogy that assumed the refraction of small revolving globules is, however, another issue. Thus, he essentially assumed, as Newton would soon do, that light consists of particles, and explained refraction (as well as reflection), as normal forces acting upon these particles at the level of the separating surface. Consequently, the tangential speed of light particles remained unchanged, that is, $V_i \sin i = V_r \sin r$, from which is derived the following expression of the law of refraction:

$$\frac{\sin i}{\sin r} = \frac{V_r}{V_i} = n\,.$$

This equation implies that, upon passing into a denser medium ($i > r$), the speed of light particles increases ($V_r > V_i$), and vice-versa, something which seems to contradict common sense, but we will return to criticism of ***the corpuscular model of light*** at a later time, when we discuss Fermat. For now, using the exact law of refraction, in chapter 8 of his book, ***La Dioptrique***, Descartes brilliantly solved the problem which Kepler failed to do in 1604, and described the surfaces (henceforth known as ***Cartesian surfaces***, or ***Cartesian ovals***) which, by definition, ensured rigorous stigmatism for pairs of points, conical sections being an important particular case of this. As we've observed at the relevant time, these surfaces found a wide field of application in the crafting of reflective telescopes and aspheric lenses. It is remarkable that Descartes suggested the use of hyperbolic lenses for perfecting the telescope, and designed a machine for grinding this type of lens, by use of which, in the same period, the optician Ferrier of Paris crafted the first ***convex hyperbolic lens***. The ***crystalline lens*** was also an object of study for Descartes, who conducted numerous anatomical dissections of the eye, and confirmed Kepler's conclusion, that the retina is the photoreceptor of images, and Scheiner's experiments, which he had conducted on cow's eyes and the human eye; moreover, he curved the crystalline of eyes prepared beforehand, stretching it by gently applying pressure, and observed that a clearer image of close objects then formed on the transparent retina, a fact which permanently elucidated ***the mechanism of sight adjustment.***

Finally, let us mention Descartes' brilliant contribution to formulation of the geometric quantitative law of the ***rainbow***
1637 (***Les Météores***, 1637), the most frequently occurring and grand natural optical phenomenon, caused by the refraction of sunlight in rain water droplets, as illustrated in figure A.2. Naturally, this phenomenon has been described as early as Antiquity (Aristotle, 350 BCE; Seneca, 50 BCE), but it was much later that Witelo (1274) attempted to explain it through refraction, and finally ***Dietrich of Freiberg***, latin ***Theodoricus Teutonicus of Vriberg*** in his work ***De luce et ejus origine, de coloribus, de iride et radialibus impressionibus***
1311 (1311) advanced the correct qualitative explanation of the rainbow that entailed sun rays refracting twice within the spherical water droplets, and reflecting once in the case of the primary (inner) arc, and twice in the case of the fainter secondary (outer) arc, since the light's intensity diminishes after each reflection. Unfortunately, this work, remarkable for the 14th century, remained hidden in a monastery, and then within the public library of Basel, undiscovered for centuries, until ***Giovani Battista Venturi*** (1746 – 1822), the Italian physician known for his treaties on hydraulics – the same one who made available some of the remaining manuscripts and scientific work of Leonardo da Vinci in 1797 – put it into circulation through his work ***Commentari sopra la storia e la teoria dell'opttica***, Bologna, 1814. Thus it was that Maurolico (1575), who also measured the angle between the incident sunrays and those that emerged – $\varphi_1 = 40° - 42°$ for the primary arc and $\varphi_2 = 50° - 53°$ for the secondary arc – failed in his attempt to explain the phenomenon, and ***Marcus Antonius de Dominis*** (1566 – 1624), in his work ***De radiis***

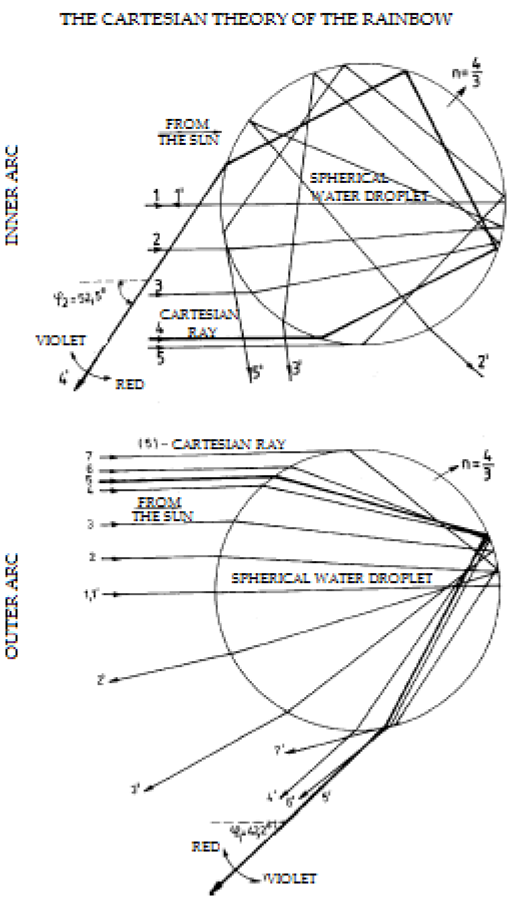


Fig. A.2 How the rainbow is produced according to Descartes

visus et lucis in perspectivis et iride, published in 1611 in Venice, demonstrated experimentally (simulating a rainbow in similar conditions, with the use of spheres filled with water and conveniently illuminated by the Sun) that the sunlight's path guessed by Dietrich of Freiberg in 1311 was valid. It is not clear whether Dominis had seen Dietrich's work, or whether Descartes was aware of his two predecessors, but Descartes (***Les Météores***, 1637) also undertook the experiment using glass spheres filled with water, and was surely the first to use the exact law of refraction to draw up the path of light illustrated in the figure above. However, beside this figure, drawn for $n = 4/3$, with just a few rays, Descartes manually calculated the paths of 10 000 (!) rays for $n = 250/187 \approx 1.3369$, and identified the angles of extreme deviation $\varphi_1 = 41.5°$ for the main arc and $\varphi_2 = 51.9°$ for the secondary arc. Obviously, it is only at angles close to these values that the emerging rays constitute parallel beams and effectively contribute to the two arcs of the rainbow, as seen by an observer located at a distance*. Beyond this, however, Descartes was not able to explain the colors of the rainbow and why exactly these colors, which form a sequence from violet (V) to red (R), are displayed in reverse order across the two arcs (see figure), but the mainstay of the theory was already firmly placed into position (the theory of white, or colorless, arcs). Newton will be the one to explain the colors of the rainbow. In his enlightening experiments using glass prisms, he discovers that white light is composed of numerous colored
1704 rays of different degrees of refrangibility (***Opticks***, London, 1704, translated into Romanian by Prof. Victor Marian with the title ***Optica***, Romanian Academy Press, Bucharest, 1970, in which Newton's theory of the rainbow may be found on pages 110 through 116). Newton's

* Descartes' problem concerning the spherical droplet can be easily solved analytically, using the method of maximum or minimum deviation. Thus, by writing i for the angle of incidence, r for the angle of refraction, and k for the number of internal instances of reflection, and referring to elementary notions of geometry, we obtain the formula $\phi = 2(i - r) + k(\pi - 2r)$ for the angle between solar and emerging rays, and $\varphi = \pi - \phi$ for the angle of observation. Using the formulas for the law of refraction $\sin i = n \sin r$ and $\cos i \ di = n \cos r \ dr$, and referring to the condition of extreme variation $d\varphi/di = 0$, we get

$$\sin i = \sqrt{\frac{n^2 - 1}{(k+1)^2 - 1}}$$

Thus, for example, for $n = 4/3$ and $k - 1$ we get $i_1 = 59.38°$, $r_1 = 40.2°$, $\phi_1 = 137.97°$, so $\varphi_1 = 42.2°$. Similarly, for $k = 2$, we get $\varphi_2 = 52.5°$.

theory is in fact merely an extension of Descartes' theory, namely its application for each color composing white light, as can be seen in the table below (calculations done by Newton):

		main arc	*secondary arc*
Violet	$n = 109/81 \cong 1.3457$	$40°17'$	$54°7'$
Red	$n = 108/81 \cong 1.3333$	$42°2'$	$50°57'$

And so it was that using the exact law of refraction Descartes and Newton explained numerically the main features of this fascinating phenomenon that is the rainbow. In this context we must also mention another name forgotten by history, ***Johannes Marcus Marci de Kronland*** (1595 – 1667), who in his work dedicated to the rainbow ***Thaumantias Iris, liber de arcu coelesti, deque colorum apparentium natura, ortu et causis***, published in Prague in 1648, he first made the connection between color and the deviations brought about through refraction, by observing the spectrum of white light (iris trigonia) produced by a triangular prism (trigonum) placed before the opening of a dark chamber. Marci's contribution was acknowledged after as much as 300 years passed from his death (the Thirty Years' War made impossible the flow of ideas between Central and Western Europe).

The issue of who first discovered the law of refraction has been much debated in literature, up to the present time. It began with Christian Huygens and Isaac Voss, both Snell's countrymen. Huygens suspected, and Voss violently accused Descartes of having previously read Snell's manuscript ("quae et nos vidimus aliquando et Cartesium vidisse accepimus, ut hinc fortasse mensuram illam quae in sinibus
1653 consistit elicuerit"*, Huygens, ***Dioptrica***, 1653), of plagiarizing Snell's law, and of purposefully making up the theory of light globules in order to mask the plagiarism (Vossius, ***De lucis natura et proprietate***, 1662). It seems strange that these assertions were made as late as three and twelve years after Descartes' death, respectively. It's true that Descartes traveled extensively throughout Europe, and that he was a permanent resident of the Netherlands for 20 years (1629 – 1649), where he lived in 13 cities, among which Leiden (in 1630), where Snell had lived. Later

* "which we have also seen and which we know Cartesius saw, and from which perhaps were taken those measures of the ratio of sines".

studies regarding this controversy (P. Kramer, ***Abhandlungen zur Geschichte der Mathematik***, No. 4, pp 233 – 278, Teubner, Leipzig, 1882; H. Boegehold, ***Keplers Gedanken über Brechungsgesetz und ihre, Einwirkung auf Snell und Descartes***, Ber. Naturwiss, Ver. Regensburg, **19**, 150 (1928 – 30); M. Herzberger, ***Optics from Euclid to Huygens***, Applied Optics, **5**, 1383 (1966)) revealed by referring to Descartes' letters that he had been preoccupied with aspheric plano-hyperbolic lenses since 1627. Since such a lens stigmatically foci the image of an object located at a distance precisely because its hyperbolic surface is defined by the law of the sine, it can be justly assumed that Descartes was aware of the law of refraction three years before his visit in Leiden in 1630. Consequently, it is fair that this law, which proved crucial in later developments in the field of optics, as well as within the general sphere of the sciences, be named ***the Snell-Descartes law***, without neglecting to mention Kepler, who had come so close to its discovery, or Witelo, Alhazen, and Ptolemy, who had had the means of discovering it much earlier.

There can be no doubt that Kepler and Descartes' laying of the theoretical foundation for dioptrics was greatly aided by the telescopes and microscopes crafted by mere chance at so late a time as the 1610s, after many centuries of polishing glass and crystals, and crafting lenses or objects similar to these. But no less true is the fact that the fundamental
1637 law of geometrical optics publicly communicated by Descartes in 1637 was a powerful impetus that allowed the scientific world to overcome its millennia-old state of stagnation. We will first witness the metamorphosis of the law of refraction into an even more general law of geometrical optics, ***Fermat's principle*** (1657), which bore profound conceptual implications with respect to the variational formulation of natural laws. Evincing the same sort of eagerness to make up for lost time, science moved on to the designing and crafting of high-performance telescopes (***Gregory, Newton, Cassegrain, Hooke, Huygens, Hadley, Dollond, Herschel***), with their well-known impact on the new post-telescopic version of astronomy and on our representations of the Universe, as well as the introduction of the microscope (***Hooke, Leeuwenhoek***), which allowed high-precision micrographic and anatomical observation. New facts were quickly accumulated, regarding subtler light-related phenomena such as diffraction (***Grimaldi***),



interference (***Hooke***), dispersion (***Marci, Newton***), birefringence (***Bartholinus***), the speed of light (***Römer, Bradley***), in a period of unprecedented intellectual fertility, crowned by Newton's and Huygens' genius. Unlike the preceding period, in which scientists had mostly done their work individually, from this time onward they begin to interact with considerably increased frequency, as the exchange of ideas proved extremely favorable to making new discoveries. The sciences finally begin to flourish, and, despite all the fluctuations and instabilities inherent to their time, the first powerful scientific societies take shape with the Royal Scientific Society of London, 1662, with its famous publication ***Philosophical Transactions of the Royal Society*** (begun in 1665), and with the Scientific Academy of Paris, 1664, with its publication ***le Journal des Savants*** (begun in 1665), followed by the academies of Bologna (1712), Berlin (1720), Petersburg (1725), Uppsala (1725), Stockholm (1739), Copenhagen (1743), and so on. Let us next outline the great moments of the post-Cartesian period.

Fermat, Pierre de (1601 – 1665), French mathematician, founder of modern number theory and early proponent of infinitesimal calculus (he developed Kepler's method, according to which variations of functions are null in the vicinity of maxima and minima). He engaged in fierce debates with Descartes, on the issue of curve maxima and minima, as well as with respect to the way in which Descartes "demonstrated" the law of refraction and reached the conclusion that the speed of light (particles) is greater in denser media. In truth, if we were to consider this controversy from a 20th century perspective, the Descartes-Newton corpuscular model of light may be easily modified and adapted to the experience which has been attained up to the present moment. For this, it is sufficient to write the conservation of the tangential component of momentum (and not of speed), i.e. $p_i \sin i = p_r \sin r$, which gives the law of refraction in its correct form

$$\frac{\sin i}{\sin r} = \frac{p_r}{p_i} = n \,,$$

which ***cannot*** be written as $\frac{\sin i}{\sin r} = \frac{V_r}{V_i} = n$, since the ***momentum*** of photons (i.e. light particles) ***cannot*** be expressed as the product $p = mV$.

As we know, (Einstein - de Broglie), $p \sim 1/\lambda = \nu/V \sim 1/V$ where ***V* *is the phase velocity*** of the monochromatic light wave. Thus, the correct expression of the law of refraction can also take the form

$$\frac{\sin i}{\sin r} = \frac{V_i}{V_r} = n \, ,$$

or, if defining the absolute index of refraction as $n_{i,r} = c/V_{i,r}$, associated with passage from vacuum to any medium,

$$\frac{\sin i}{\sin r} = \frac{n_r}{n_i} = n \, ,$$

where n (simple) is the relative index of refraction, associated with passage from the considered medium (i) to medium (r). But all these representations were not this clear at the time of the search for concepts during the famous Descartes versus Fermat dispute, which was taken up again between Newton and Huygens, and then, reaching a scale of hundreds of years, between the 18th and 19th centuries, until the 20th century reconcilement of (or postponement of debate regarding) the wave-particle duality. We may ask why an experiment was never conducted from the start to decide if, for example, light traveled at a greater speed in water than in air, as it could be concluded from Descartes' original formula, or if it was the other way around, as Fermat would have it. The answer would be that the experimental techniques available throughout the 17th and 18th centuries were not suitable for measuring the variation of the speed of light in traversing denser media. It was only possible to determine the speed of light in vacuum, using
1676 astronomical methods (***Römer***, 1676, and ***Bradley***, 1727), as we shall
1727 soon see. It was not until the 19th century (the century of wave optics), that it became possible to prove experimentally, under terrestrial conditions, that the speed of light is greater in the air than in denser media, such as water (***Fizeau*** and ***Foucault***, 1850). But at that time the wave model of light had already been consolidated through the demonstrating and elucidating of its distinctive phenomena of interference (***Young***'s experiments, 1802 – 1804), diffraction (***Fresnel***'s theory and experiments, 1818), and polarization (the transverse nature



of light waves, ***Young*** and ***Fresnel***, 1820), advancing with great strides towards the epitome of its glory, the electromagnetic theory of light (***Maxwell***, 1873 and ***Hertz***, 1887). In the 17th century, however, it would have been risky to advance any hypothesis regarding the subtle nature of light based solely on the observational and experimental data available at that time. Of course, it has sometimes happened that certain statements have subsequently been confirmed, as is the case with Descartes' suggestion that light curves upon passing massive celestial bodies, which anticipated Einstein's conclusions by close to 300 years*. However, the only sound basis on which to further the knowledge of luminous phenomena was the exact mathematical expression of the law of refraction, whose experimental confirmation in the media of glass, crystals, water, and many other liquids had become a common endeavor after the publication of Descartes' work on optics. At first, Fermat was convinced that the law itself was false, but experimental evidence soon became so unequivocal, that the only thing left to do was to unravel the mystery of Descartes' corpuscular model, based on which he had indeed reached the formally exact law, which, however, yielded the paradoxical conclusion that light encountered less resistance in a denser medium than in a more rarefied one. After many years of searching, Fermat eventually recalled Hero's principle of the shortest path of light (50 CE), and managed to eliminate the speed paradox, formulating a great postulate (see his posthumously published works, ***Varia opera mathematica***, p. 156, Toulouse, 1679), known as ***the principle of least
1657 time***, or ***Fermat's principle*** (1657). First expressed in its Heronian form, "… que la nature agit toujours par les voies les plus courtes" (in a Toulouse letter, August 1657), and then in its general form, that also encompasses refraction, namely that ***the path taken between two points by a ray of light is the path that can be traversed in the least time*** (in a Toulouse letter, 1st of January 1662), with the later addition that ***the interval in which light propagates is stationary*** (minimal, maximal, or constant), this principle constitutes the most concise law of geometrical optics. From it first follow all the laws pertaining to refraction, reflection,

* What's more, today's astronomers have discovered veritable gravitational mirages ("double quasars" and "multiple quasars"), that they explain are produced through the deformation, amplification and multiplication of far-off bodies, effected by the mass of galaxies located close to the trajectory of light rays (see Alain Blanchard, ***Les mirages gravitationnels***, La Recherche, 1987).



and linear propagation in homogenous media that had been formulated up to that point in time. Moreover, by way of addition of infinitesimal intervals of time, it becomes much more general and may be applied to curvilinear paths along any non-homogenous medium. Thus, for example, in a case of refraction at a surface separating two homogenous media, we have

$$t_i + t_r = \frac{s_i}{v_i} + \frac{s_r}{v_r} = \text{stationary},$$

from which the same law of the ratio of sines may be deduced, but in its physically acceptable form $\sin i \,/ \sin r = v_i/v_r = n$. In its more general form, which would also apply to non-homogenous media, Fermat's principle is written as

$$\int_{t_1}^{t_2} dt = \int_{p_1}^{p_2} \frac{ds}{v} = \text{stationary}.$$

The conceptual leap that Fermat achieved is evident, from a sine law that directly imitates the geometry of refraction (and which by itself constitutes a simple rule that replaces interminable goniometrical tables), to a single principle that governs the behavior of light rays in any situation. Fermat's concept of building up the edifice of natural law based on an integral principle of extrema greatly inspired his renowned successors Leibniz, Jean and Jacques Bernoulli, Maupertuis, Euler, Lagrange, Hamilton, Jacobi, up to Feynman in the second half of the 20th century.

So it was that ***Gottfried, Wilhelm Leibniz*** (1646 – 1716) again took up the issue in the Aristotelian spirit that „nature does nothing in vain," and advanced ***the principle of least effort*** (***Acta Eruditorum***, Leipzig,
1682 1682), where by effort he meant mechanical work, namely the product between the force of resistance R encountered by light in a medium and the distance traversed s. By applying infinitesimal calculus (which he had developed in a race against Newton) to his principle $R_i s_i + R_r s_r =$ minimal, he reached the expression $R_i ds_i + R_r ds_r = 0$, from which, by expressing the path-related elements using angles, could be deduced the



law of sine ratio in the form $\sin i\,/\sin r = R_r/R_i$, which meant that resistance was greater in denser media. So nothing essentially new compared to Fermat's results if we simply base our judgment on the hypothesis that $v \sim 1/R$. But, in order to illustrate how difficult and slippery the way of physical concepts and representations is, let us recall that Leibniz made the ***incredible*** supposition that $v \sim R$ (based on an analogy to the flow of a river being greater at the narrower points of its course). Leibniz's surprising choice may be justified especially if we consider the trend of thinking dominant in his time, imposed by the authority of Descartes and Newton, which stated that light traveled at greater speeds in denser media, but also if we take into account the fact that the generous formalism introduced by Fermat in geometrical optics would momentarily be taken up in its kindred fields of study, in mathematics and mechanics. For ***Jean Bernoulli*** (1667 – 1748) would formulate his famous ***brachistochrone problem*** (***Acta Euroditorum***),
1696 Leipzig, 1696), to which he would also give the solution by making an analogy to geometrical optics, which his older brother ***Jacques Bernoulli*** (1654 – 1705) would go further and reformulate for a wider category of problems, thus laying the foundations of ***the calculus of variations***. ***Pierre Louis Moreau de Maupertuis*** (1698 – 1759) similarly took up Leibniz's idea and formulated ***the principle of least action*** (***Mémoires de***
1744 ***l'Académie de Paris***, 1740, 1744), where by ***action*** he meant the product between a particle's momentum mV (not necessarily a particle of light) and the distance traversed s. This new principle may be expressed as $V_i s_i + V_r s_r =$ minimal, and, as we have seen, yields the Cartesian law in its primitive form $\sin i\,/\sin r = V_r/V_i$. But the great discovery was that, although this principle did not apply to light rays in geometrical optics, it was suitable for trajectories of mechanical motion. Actually, the precise expression of Maupertuis's principle, applicable to a particle's motion in a conservative field, was given by ***Leonhard Euler*** (1707 – 1783) in ***Methodus inveniendi lineas curvas, maximi minimive proprietate gaudentesi***, Lausanne et Genève, 1744:

$$\int_{p_1}^{p_2} V ds = \text{stationary},$$

and for more general conditions, by ***Joseph Louis Lagrange*** (1736 – 1813) in ***Mécanique Analytique***, 1788, by ***William Rowan Hamilton*** (1805 – 1865) in ***On a General Method in Dynamics***, Philosophical Transactions of the Royal Society, 1834, by ***Karl Gustav Jacob Jacobi*** (1804 – 1851) in ***Vorlesungen über Dynamik***, 1866, for classical mechanics, up to ***Richard Phillips Feynman*** (1918 – 1988), in ***Space-Time Approach to Non-Relativistic Quantum Mechanics***, Reviews of Modern Physics, **20**, 367 (1948), and in ***Quantum Mechanics and Path Integrals***, 1966, for quantum mechanics. And so, the elegant variational formulations of natural laws and the associated Lagrangian and Hamiltonian formalism dominated the physics of 19th and 20th centuries. Such was the conceptual enlargement that sprang from Fermat's principle. Of course, we may add post festum that if Hero had extended his principle of the minimum to refraction, then it would have been he who authored what is now known as Fermat's principle, but, like Ptolemy faced with the law of the sines, he missed his great chance of jumping past thousands of years' time.

Let us return to practical applications and observe that the achievements of dioptrics first unlocked the practical possibilities of catoptrics. Thus, the main flaw of the first telescopes, namely the chromatic aberrations (the image's iridescence), as well as the degree of perfection that had been reached with metallic concave mirrors, used as early as Antiquity (see the legends concerning Archimedes' mirrors or the lighthouse of Alexandria, or even the supposed mirror instrument of Ragusa), led to the concept of the ***reflective telescope***, first proposed in a form using spherical reflector by ***Nicolaus Succhius*** (***Optica philosophica***, 1616), then in a parabolic reflector version by ***Marin Mersenne*** (***Cogitata phisico-mathematica***, 1644), but effectively designed by the Scottish mathematician ***James Gregory*** (1638 – 1675) in
1663 his memoir ***Optica promota*** (1663), in which he describes the version that bears his name (using a parabolic main mirror and an elliptical concave secondary mirror). In turn, Gregory's work inspired ***Robert Hooke*** (1635 – 1703), who effectively built Gregory's telescope, but with the use of spherical mirrors (Philosophical Transactions of the Royal Society, 1674), ***Giovanni Cassegrain*** (1625 – 1712), who described his shorter version (using a hyperbolic convex secondary mirror) in ***Journal des Savants*** (1672), and the then-young ***Isaac Newton*** (1642 – 1727). As

early as 1666, Newton had discovered the phenomenon of the dispersion of light, or, in his terminology, the phenomenon of different refringence of light rays of various colors, based on which he had concluded that it was impossible to design perfect telescopes, that is, telescopes that would show clear, non-iridescent images in white light. This conclusion, known as "Newton's error," (see Chapter II, section 2.7), based on the assumption that the relative dispersion of light is the same in all glasses, caused Newton to believe that the problem of creating an achromatic lens telescope was insoluble. But even this error came to a happy conclusion when Newton set out to build, with his own hands and a level of skill that may be envied even today, the reflective telescope that bears his name, achieving a visual augmentation of 30 – 40X, using an instrument only 6 inches long and 1 inch in diameter (***Phil. Trans. Roy.***
1672 ***Soc.***, 1672). This history of Newton's telescope, as well as the entirety of his invaluable contributions to the experimental research concerning optical phenomena, are worth studying directly in their author's works (see Isaac Newton, ***Optica***, Romanian Academy Press, Bucharest, 1970). The Royal Society of London bears the honor of having appreciated Newton to his full merits. He was elected member during the session of the 11th of January 1672, and his telescope was exhibited in the Society's Library, where it can be found even today, alongside the inscription "Invented by Sir Isaac Newton and made with his own hands, 1671." But the first reflective telescopes, designed by Hooke and Newton, were small experimental models, and their value lay only in the path towards greater performance along which they pointed. The first such larger telescopes which proved to be of practical use were fashioned as much as 50 years later, in Newton's design (in 1723), as well as following Gregory's model (in 1726), by ***John Hadley*** (1682 – 1744). But the advantage mirrors held over lens (the elimination of chromatic aberrations and far greater magnifying power) further spurred the search for yet greater performance. So it was that the most powerful telescopes of the following period were built by the famous English astronomer ***Sir William Herschel*** (1738 – 1822). Among them, his great telescope with a focal distance of 12 meters (***Phil. Trans. Roy. Soc.***, p.
1795 347, 1795), built following a personal design, which bears his name, and is still used today. With the help of these wonderful devices, Herschel discovered the planet Uranus (thus doubling the extent of the known



solar system), as well as over 800 double stars (referring to which he demonstrated that binary stars revolve around each other in accordance with Newton's law of gravity), he observed and catalogued around 2,500 nebulae, or clusters of stars, and made the sensational discovery of the "invisible light" of the Sun's infrared spectrum (***Phil. Trans. Roy. Soc.***,
1800 p. 292, 1800; by exposing thermometers to light from different areas of the solar spectrum, he observed that the temperature indicated became greater as he moved the thermometer from violet to red, and reached its maximal value within the invisible area, known today as infrared).

It is remarkable that determining ***the speed of light*** (in vacuum), one of the greatest achievements in the field of optics, was first possible using the astronomical methods developed in the 17th century (Römer,
1676 1676, and Bradley, 1727). Until the year 1676, it was not known whether or not light travelled instantaneously, and, if it didn't, how great its speed might be. Beginning with Antiquity, the dominant view became that of Hero (***Catoptrica***, around 50 CE), who believed, on the grounds of analogy to the trajectory of an arrow shot with increasingly greater speed, that the speed of light, which travels on a linear path, must be greater than the speed of any given terrestrial body. Descartes (***La Dioptrique***, 1637) believed that, if the speed of light were finite, then, by interacting with the orbital speed of the Earth, we should be able to observe an apparent movement of the "fixed stars" across the celestial sphere (this effect, termed stellar aberration, was indeed discovered by Bradley in 1727, after 90 years). Galilei's attempt (1638), using light signals exchanged between two terrestrial observers located at a distance of around 3 km, failed, naturally, because it took a very short time for light to travel between them. And so, we arrive at the Cassini – Römer moment (1675 – 1676), at the moment of Italian-born astronomer ***Giovanni Domenico Cassini*** (1625 – 1712), member of the French Academy and of the Royal Society, the first director of the Astronomical Observatory of Paris, which he administered for 40 years (1671 – 1711), with a prodigious career (around 200 scientific advancements, among which, the discovery of Saturn's 4 satellites and the study of the planets' rotation around their own axes) and of Danish astronomer ***Olaf Christensen Römer*** (1644 – 1710), himself a member of the French Academy, who later (1681 – 1710) would become professor at the University of Copenhagen, head of its Astronomical Observatory, and

1675 even the city's mayor (1705). In August of 1675, Cassini announced "la seconde inégalité" in the motion of Jupiter's satellites, an effect which he wrote "seems to be caused by the fact that light requires a certain interval of time to reach us from the satellites, and that it takes 10 – 11 minutes for it to traverse a distance equal to the radius of the Earth's orbit."

As has been known since Galilei's time (1610), Jupiter's satellites complete their revolution in several days (for example, its first satellite, ***Io***, the one closest to the planet, completes its revolution in 42 hours, 28 minutes, and 16 seconds, and its orbit around Jupiter is virtually coplanar with Jupiter's orbit around the Sun). Cassini's observed irregularity consisted of the fact that the time between two successive eclipses of Jupiter's ***satellites*** increases and decreases according to the periodical increase and decrease, respectively, of the distance between Jupiter and the Earth, as they follow their respective orbital paths around the Sun (see figure A.3). Römer and Cassini's observations in the year 1676 showed that Jupiter's first satellite emerged from the planet's shadow (it "arose") 10 minutes later in November than in August, that is, after a quarter year, in which time the distance between the Earth and Jupiter practically increased by the length of the terrestrial orbit radius, thus confirming the conclusion Cassini had reached the preceding year. Based on this delay, caused by the propagation of light, and on the data then available regarding the radius of Earth's orbit, Römer calculated the speed of light at around 214,000 km/s, a truly large value, but still ***finite*** (the 10-minute delay was overestimated; the current data show that light traverses the 149.10^6 km distance between Sun and Earth in 8.35 minutes, which means that $c \approx 300{,}000$ km/s). It would be interesting to note that Descartes' argument regarding stellar aberrations, formulated almost 40 years prior, had so strong an influence, that Cassini, the first proponent and co-author of the first method used to calculate the speed

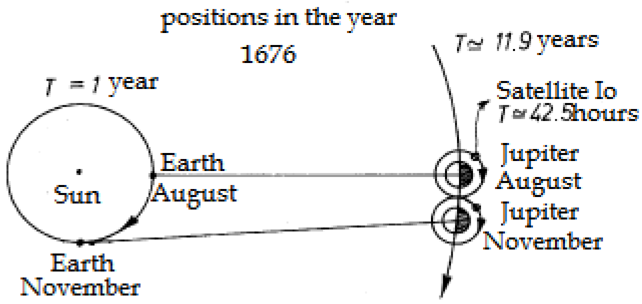


Fig.A.3. Determining the speed of light (Römer method)



of light, denounced this result, and that Römer was left to announce it alone, to a general reaction of fierce hostility, during the Science Academy assembly on the 21st of November 1676, and then publish it in the ***Journal des Savants*** edition of 7 December 1676 and in the 1677 ***Philosophical Transaction of the Royal Society***. This discovery was, however, not recognized by the scientific community of the time until
1727 the Cartesian objection was satisfied in 1727, when British astronomer ***James Bradley*** (1693 – 1762), professor at Oxford University and member of three academies (the Royal Society of London, the French Academy of Sciences in Paris, and the German Academy of Sciences at Berlin), finally discovered the phenomenon of stellar aberrations (***Phil. Trans. Roy. Soc.***, 35, p. 637, 1728). By carefully observing the positions of stars γ and δ of the Draco constellation, located virtually at the pole of the Earth's ecliptic, Bradley discovered that they indeed each describe a small ellipsis over one year's time, with an angular diameter of 2α ≈ 40". By analogy to the necessary inclination of an umbrella during rain (see Fig. A.4), the effect of stellar aberration would mean the observation telescope is inclined at 20" $= \alpha \approx 20'' = 10^{-4}$ radians. And so, Bradley obtained:

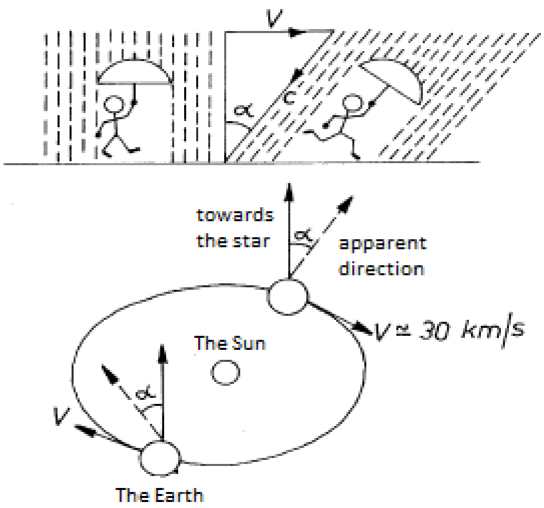


Fig. A.4 To understand Bradley's aberration

$$c \cong \frac{V}{\alpha} \cong \frac{30}{10^{-4}} km/s = 300{,}000 km/s$$

thus confirming Römer's results, and decidedly determining the value of the first universal constant. More than 100 years would pass before ***Armand Hippolyte Louis Fizeau***, with his cogged wheel (***C. R. Acad. Sci.***, Paris, 1849), and ***Jean Bernard Léon Foucault***, with his rotating mirror (***C. R. Acad. Sci.***, Paris, 1850), would inaugurate a series of terrestrial measurements of *c*, whose currently accepted value (determined in 1970)

is

$$c = 299{,}792.458\ km/s,$$

with an error of 300 m/s at most, that is, a 0.0001% error. A considerable number of experiments have been undertaken to more precisely determine the exact speed of light. Because of its high value, the use of great distances or small time intervals is necessary. Alternatively, one can measure the λ_0 wavelength in a vacuum of an electromagnetic wave of known frequency ν, and calculate the speed of light using the simple formula $c = \nu\lambda_0$. In fact, the entire arsenal of classical and quantum electrical devices of the 20th century have been used to calculate this value, as a sort of performance test, because of its importance in the whole evolution of physical theories. Is light "fast"? We may get a sense of its speed by simply considering the fact that it takes 8 minutes for light from the Sun to reach us, as we have seen, that it reaches us from the next closest stars (Alfa Centauri and Proxima Centauri) in 4.3 years, from the most brilliant star (Sirius), in 8.5 years, it traverses the distance between the Polar Star and Earth in 400 years, that between the Earth and the center of our Galaxy, in 5×10^4 years, it reaches us from the closest galaxies in 2.5×10^6 years, from the farthest observed galaxies, in 6.5×10^9 years, and, finally, from the edges of the known Universe, in around 1.9×10^9 years. It is an essential feature of the structure of our Universe that the speed of light in a vacuum cannot be matched by any other material body, that it cannot be surpassed by any given signal, that light travelling through a vacuum is the fastest medium of information transmission, and that ***c constitutes a speed limit***, a truly significant assertion, raised by ***Albert Einstein*** to the rank of postulate of the theory of relativity. Another remarkable feature of its subtle nature is the fact that ***light also propagates through void space***, that is, through space devoid of the usual atomic and molecular matter (unlike sound, for example, whose propagation depends on its presence), a phenomenon which suggests that this "void" must be the setting of material properties and processes of a most refined and profound nature.

Whereas the telescope has been used to discover the Universe of bodies far away, the microscope has allowed the discovery of the Universe of infinitesimal bodies. ***Robert Hooke*** (1653 – 1703), mentioned

earlier for building the first reflection telescope (1674) and significantly improving the quality of lenses, also built the first practically suited
1675 ***compound microscope*** (1675), thus pioneering the field of precise microscopic observations. These observations were presented in his ***Micrographia or some physiological descriptions of minute bodies***, London, 1665, which soon became a milestone work in the history of the microscope. Hooke played a key role in instituting the ***Royal Society of London***, as a founding member (1663), curator by office for life of the Society's experiments and weekly sessions (1662 – 1703), and one of the Society's secretaries (from 1677 onwards). Among the first elected as ***Fellows of the Royal Societies (F.R.S.)*** were illustrious names such as ***Huygens*** (1663); ***Newton*** (1672); ***Flamsteed*** (1676), the first director of the Greenwich Astronomical Observatory (1675); ***Halley*** (1678), renowned for having calculated the orbit of the eponymous comet, 1682, the author of the famous formula $(1/p_1) + (1/p_2) = (n - 1) \times [(1/r_1) - (1/r_2)]$, and, finally, one of Hooke's famous rivals; Dutch biologist ***Antony van Leeuwenhoek*** (1632 – 1723), F.R.S. from 1679 onwards, who astounded his contemporaries with his discoveries made using the ***simple microscope*** (***Phil. Trans. Roy. Soc.***, 1673), author of the four-volume work ***Arcana naturae ope microscopiorum detecta*** (***The Secrets of Nature Revealed through a Microscope***), Leiden, 1722. Using lenses fashioned by his own hand with great precision, the smallest of which having a diameter of only 0.5 mm, and samples fixed using needles, Leeuwenhoek was able to observe biological structures and processes of a kind never before seen, achieving detailed resolutions close to the theoretical limit of one micron. With the help of such an instrument, he was able to study the texture of tooth and bone, striated muscle tissue, the fibrous crystalline lens, the circulation of blood through the capillaries (thus confirming and elaborating on microscopic observations previously made by Italian physician ***Marcello Malpighi***, 1661), he gave the first exact description of red blood cells, bacteria, and protozoa, and was the first to observe the permanent chaotic motion of infinitesimal particles suspended in liquids (1673). This ***Brownian motion*** was confirmed in 1827 by Scottish physicist ***Robert Brown*** (1773 – 1858), who used a microscope to observe the continuous zig-zag motion of pollen granules suspended in a drop of water. The kinetic theory of Brownian motion in suspensions, caused by molecular

collisions, was advanced by ***Albert Einstein*** (1905) and brilliantly confirmed through experimental means by ***Jean Perrin*** (1908), in his classical microscopic study of smoke particles in the air (diameter $\cong 10^{-4}$ cm), following which a useful value for ***Avogardo's number*** was determined.

From the 17th century onwards, the study of refraction gathered great momentum due to the practical and scientific interest for building optical instruments. The shapes of lenses and mirrors, the geometrical construction of images, an introduction to the study of aberrations: all these were the works of such figures as ***Johannes Kepler*** (***Dioptrice***, 1611), ***Francesco Bonaventura Cavalieri*** (***Exercitationes geometricae***, 1647), ***Christiaan Huygens*** (***Dioptrica***, 1653), ***Isaac Newton*** (***Lectiones opticae***, 1669), ***Isaac Barrow*** (***Lectiones opticae et geometricae***, 1674), and ***Edmund Halley*** (***Phil. Trans. Roy. Soc.*** , 1693), which were systematically ordered in the great optical treatises of the 18th century, written by ***David Gregory*** (***Optics***, 1735), ***Robert Smith*** (***A Compleat System of Opticks***, four volumes, 1738), ***Joseph Harris*** (***A Treatise of Optics***, 1775), and were popularized through the widely accessible works of ***James Ferguson*** (***Astronomy explained upon Sir Isaac Newton's Principles, and made easy for those who have not studied Mathematics***, 1756; ***Lectures on Selected Subjects in Mechanics, Hydrostatics, Pneumatics and Optics***, 1760). Considerable effort was undertaken in the years 1655 – 1660 for the perfection of refraction telescopes by the brothers ***Christiaan*** and ***Constantin Huygens***, and by the Italians ***Eustachio de Divini***, ***Giuseppe Campani***, and the great pupil of Galilei, ***Evangelista Torricelli*** (the first demonstration that a small glass sphere, as could be easily obtained in a fire, was the most powerful magnifying glass; thus, an explanation was discovered for why plants should not be watered under intense sunlight: the spherical droplets of water concentrate solar energy in focal points close to the surfaces of leaves, which they scorch locally). They fashioned exquisitely cut and polished telescope objectives with focal distances of between 30 and 40 meters, but the iridescence caused by chromatic deviations that had been so carefully studied by Newton in his experiments with lenses and prisms (***Lectiones Opticae***, 1669; his dissertation ***A New Theory about Light and Colours***, presented during the Royal Society meeting on the 6th of February 1672; his ***Opticks*** treatise of 1704) made obtaining images

of satisfactory quality impossible. Newton's famous "error" regarding the essential impossibility of a lens to refract light composed of different
1758 colors into a single focal point was ultimately solved in 1758 by the London-based optical physicist ***John Dollond*** (1706 – 1761), who introduced the ***achromatic doublet***, obtained by attaching together a crown glass convex and a flint glass concave lens. For his remarkable achievement in perfecting refraction optical instruments, Dollond received an inventor's patent (his descendants and the Dollond and Aitchison company in London are active to this day), and was given the honor of being elected F.R.S. (he presented his work, ***An Account of some Experiments concerning the different Refrangibility of Light***, before the Society in 1758). But, like all discoveries, the invention of the achromatic doublet has a context and a history of its own. In this instance, it had actually been discovered in 1733 by the optical science amateur and professional lawyer ***Chester Moor Hall*** (1703 – 1771), who managed to practice his art in secret for 25 years, until Dollond happened to learn of it during a visit to the shop of lens polisher George Bass. However, the discovery found Dollond well prepared, since he had been long interested in the issue of achromatic lenses, and had been in correspondence with the great mathematician ***Leonhard Euler***, who was also preoccupied with the theory of colors and the compensation of chromatic aberrations using optical media of opposite dispersion (***Nova
1746 theoria lucis et colorum***, Mem. Acad. Berlin, 1746) and with ***Samuel Klingenstierna*** (1689 – 1785), professor of mathematics and physics at Uppsala University, member of the Academy of Sciences in Paris and of the Royal Society of London, who was inspired by Euler to discover Newton's error and developed a mathematical theory of the achromatic
1760 objective (published in 1760). History thus honors Hall with the title of inventor, but also Euler and Klingenstierna for their fundamental research, and Dollond for introducing this highly consequential invention into common use. Achromatic lenses preoccupied Euler for many years. He effectively designed and built an eight-lens telescope, and calculated the effectiveness of numerous lens combinations (***Dioptrica***, three volumes, Petersburg, 1769 – 1771). Still, the study of light dispersion by means of spectral separation with the use of prisms that Newton had so elegantly inaugurated in the 1670s was not continued until 1800. It was then that ***William Herschel*** (***Experiments on***

the Refrangibility of the Invisible Rays of the Sun, Phil. Trans. Roy. Soc.
1800 1800), discovered, as we have seen, the existence of invisible ***thermal radiation*** (***infrared***), which is reflected and refracted according to the Snell-Descartes law. ***Johann Wilhelm Ritter*** (1776 – 1810) repeated Herschel's experiment, but observed the blackening of silver chloride as a detector, and discovered the existence of ultraviolet radiation, by observing that the effect of ***chemical radiation*** (***ultraviolet***) is weaker in red, but reaches a maximal value just beyond the violet end of the visible
1801 spectrum (***Gilberts Annalen***, 7, 525, 1801). It would also be interesting to
1826 note that the first ***durable photograph*** was taken as late as 1826 by ***Joseph Nicéphore Niépce*** (1765 – 1833), who used a camera obscura with convergent lens and a photosensitive plate fashioned from a copper alloy, which he exposed for around eight hours; the (negative / positive) photographic process familiar today, entailing the capturing of images on paper covered with silver chloride fixed with sodium chloride, was introduced in 1835 by ***William Henry Fox Talbot*** (1880 – 1877).

1802 Let us now return to the year 1802, in which ***William Hyde Wollaston*** (1766 – 1828) made two remarkable observations (whose significance was to be understood by Fraunhofer fifteen years later). In short, Wollaston repeated the simplest of prism experiments, described by Newton as follows: in a very dark room, he placed a glass prism before a round aperture in the blinds of a window, about a third of an inch in diameter, through which a solar beam might be refracted mainly upon the opposite wall of the room, and there form a colored image of the Sun (see Isaac Newton, ***Optica***, p. 26, translated into Romanian by prof. Victor Marian, Academy Press, Bucharest, 1970). Wollaston's version (***A method of examining refractive and dispersive powers by prismatic reflection***, Phil. Trans. Roy. Soc. II, pp 365-380, 1802) was different in that he viewed ***a narrow slit***, strongly illuminated by the Sun, ***directly through the prism***, and so observed for the first time dark lines cutting across the solar spectrum (the Fraunhofer absorption lines, of course). Wollaston then used candlelight as a light source and made the first observation of several bright lines against a dark background (that is, the atomic emission spectral lines, among which, naturally, the unsolved sodium D doublet). Like Newton 130 years before him, Wollaston actually found himself in the chamber of a large and rudimentary prism spectroscope, but did not realize that this was

prelude to the birth of quantum physics (which was close at hand at this moment in the millennial history of optics). The final step towards the materialization of the idea behind Newton's famous experiment in the form of the ***modern spectroscope*** was made by the experimenter of genius ***Joseph von Fraunhofer*** (1787 – 1826). Forced from an early age to earn a living off fine optics and mechanics, he was at the time studying telescope lens achromatization, after having attempted to achromatize combinations of prisms fashioned from theretofore tested types of glasses. In order to obtain precise measurements of the refrangibility of variously colored light beams, he used the now familiar optical system comprising a parallel beam of light limited by a narrow slit, incident on a glass prism fixed in a position of minimal deviation, which he observed through a telescope. It was using this remarkable instrument that Fraunhofer discovered (***Bestimmung des Brechungs und Farbzerstreungs Vermögens verschiedener Glassorten in Bezug auf die Verrolkommung achromatischer Fernröhre***, Denkschr. der Münch. Akad. d. Wiss, 5, 193,
1817 1817) that the continuous spectrum of sunlight is cut across (graduated sui generis) by endless, fine, more or less dark spectral lines, an ideal marker for positioning colors in the spectrum, which allowed a rigorous description of the dispersion of optical media (for example, see the table in section 2.7, ***Chromatic Aberrations***), and the selection of the most appropriate types of glasses of building achromatic optical systems. These lines have since been termed ***Fraunhofer lines***, of which the most intense (the darkest within the Sun's spectrum) are marked using letters of the Latin alphabet. Now that the dispersion and refraction indices of various types of glasses were precisely known, it was finally possible to construct the great refraction telescopes. Fraunhofer's second monumental achievement was the invention of the diffraction grating* (***Neue Modifikation des Lichtes durch gegenseitige Einwirkung und Beugung der Strahlen und Gesetze derselben***, Denkschrift der K. Akademie zu München, 8, 1, 1821-22), a much more dispersive element that the optical prism, which also allowed the precise description of

---

* In truth, the diffraction of light into multiple spectral orders by the use of a matrix of parallel wires had been observed as early as 1785, by American astronomer ***David Rittenhouse***, but his simple observations as such did not stir interest and were quickly forgotten. It was ***Thomas Young*** who through his famous experiments of 1801 – 1803 (see ***Lectures on Natural Philosophy***, London, 1807) finally demonstrated the interference of light waves originating from coherent sources and elaborated the process through which the wavelengths for different colors might be determined.



colors and spectral lines using wavelengths (the Fraunhofer theory regarding diffraction through optical gratings was finalized in 1835 by ***Friedrich Magnus Schwerd*** in his comprehensive work ***Die Beugungserscheinungen aus den Fundamentalgesetzen der Undulationstheorie analytisch entwickelt***).

In his first experiments, he used gratings of thin parallel metal wires taught at equal intervals by use of two parallel screws, the grating period being equal to the lead of the screw (gratings of around ten wires per millimeter). Dissatisfied, Fraunhofer then fashioned much denser gratings, of up to 300 lines per millimeter, by tracing equidistant lines on a glass surface, using a dividing engine with a diamond cutting tool (a process which is still used today). With the help of this grating, he extended his measurements to the dark line spectrum of direct sunlight, or of sunlight reflected by the Moon or Venus, thus observing up to 576 of these "Fraunhofer lines", as well as to the bright line spectrum of terrestrial light sources (fire, sparks and electric arcs). Of particular importance was his observation of the fact that the dark D line of the Sun's spectrum coincides with the bright yellow line of sodium. And so, during the short life he was given (he died at no more than 39 years old), Fraunhofer introduced the fundamental instruments of modern optical spectroscopy. As homage paid to the man who designed the first high precision achromatic refraction telescope and proved that the spectral lines originating from astral elements are the same as those of terrestrial source, his grave in ***München*** bears the epitaph "***Approximavit Sidera***", "***He brought the stars closer to us***".

An important contribution to fashioning achromatic objectives of large diameters (and, consequently, of high resolving power, see equation (249)) was the perfection of the manufacturing process of ***crown*** and ***flint*** glass by the Swiss optician ***Pierre Louis Guinand*** (1748 – 1824), who used an agitator to ensure the elimination of gas bubbles and tensions, and the homogenization of optical glass paste during cooling. Guinand's glass had the following compositions: crown glass, weakly dispersive (72% SiO_2, 18% K_2CO_3, and 10% CaO), flint glass, strongly dispersive and refractive (45% SiO_2, 12% K_2CO_3, and 43% PbO) and was much purer, more transparent and more homogenous (no striae) than before. Its production was soon monopolized by the large companies ***Feil*** in Paris and ***Chance*** in Birmingham. This was the glass

used to build the great refraction telescopes, from with the famous telescope with an objective 24 cm in diameter, built by Fraunhofer (first Guinand's apprentice, then his associate) and installed in Tartu, Estonia, in 1824, at the request of astronomer ***Friedrich Georg Wilhelm von Struve***, to the refraction telescopes with huge objectives, designed by the American optician ***Alvan Clark*** (1804 – 1887) and installed in the great astronomical observatories, among which the one in Pulkovo, Russia ($D = 75cm$, in 1885), one in Lick, U.S.A. ($D = 91cm$, in 1888), and one in Yerkes, U.S.A. ($D = 102cm$, in 1897). The production of high quality optical glass generally gave great impetus to instrumental optics (the telescope, the photographic objective, the spectral instruments).

Let us first follow the progress of spectroscopy, whose experimental foundation was laid by Fraunhofer, as we have seen. The prism spectroscope was subsequently perfected by ***Meyerstein*** (1856, 1861), ***Amici*** (1860) introduced the system of direct vision prisms (entailing the alternative attachment of a flint prism between two crown prisms, or of two flint prisms between three crown prisms, so that the median wavelength of the visible spectrum appears undeviated), ***Geissler*** (1856) invented the low pressure gas discharge tubes, thus supplying a new light source for emission spectroscopy. Spectroscopy is soon to become the finest and most precise method of investigating the intimate processes of light emission and absorption that take place in atomic and molecular systems, triggering a new revolution in the study of the Universe and of the structure of matter. Thus, to cite only some of the plethora of illustrious names, the great theoretician ***Gustav Robert Kirchhoff*** (1824 – 1887) and the astute experimenter ***Robert Wilhelm Bunsen*** (1811 – 1899), in the scientific environment of university town Heidelberg, lay the foundations of ***spectral analysis***, an ultrasensitive method of determining the chemical composition of terrestrial and cosmic substances (***Chemische Analyse durch Spektralbeobachtungen***,
1860 Poggendorff Annalen, 1860; ***Untersuchungen über das Sonnenspektrum und Spektren der chemischen Elemente***, Abhandl, Berlin, Akad.,
1863 1861 – 1863). Of special notice is the activity of Kirchhoff, who created a mathematical basis to the scalar theory of light diffraction, explained the Fraunhofer lines as absorption lines in the cooler gases of the solar atmosphere, and advanced his famous law according to which ***the ratio between emission power and absorption power of bodies is a universal***

function of frequency and temperature. In Sweden, ***Anders Jonas Ångström*** (1814 – 1874), a classical figure associated with high precision spectroscopy measurements (the unit of length 1Å = $10^{-10}m$ bears his name), determined the absolute wavelength of 1,000 Fraunhofer lies in the Sun's spectrum (***Recherches sur le spectre normal du Soleil***, Uppsala,
1868 1868). ***Henry Augustus Rowland*** (1848 – 1901) invented the concave grating and perfected the technique of tracing high power diffraction optical gratings of up to 1,700 lines per millimeter, thus opening the way for high resolution spectroscopy of frequencies from infrared to visible, to vacuum ultraviolet (***Manufacture and Theory of Gratings for Optical***
1882 ***Purposes***, Phil. Mag., 13, 469, 1882; Table of the Solar Spectrum Wavelengths, Astrophys. Jour., 1-6, 1895 – 1898). ***Johann Jakob Balmer*** (1825 – 1898) formulated the renowned law of wavelengths pertaining to the lines of the visible spectrum of hydrogen (***Notiz über die Spectrallinien des Wasserstoffs***, Wied. Ann., 25, 80, 1885). ***Johannes Robert Rydberg*** (1854 – 1919) discovered the general law of the frequencies of lines pertaining to the spectral series of hydrogen, namely $\nu = R[(1/m^2) - (1/n^2)]$, in which appears the universal constant $R = 3.2869\times10^{15}\ sec^{-1}$, that bears his name (***Recherches sur la constitution des spectres d'emission des éléments chimiques***, Kongl. Svenska
1890 Vetensk. Akad. Handling, 23, 155, 1890). The next step was taken by ***Walter Ritz*** (1878 – 1909), with the ***combination principle***, according to which any spectral line's frequency may the expressed as $\nu = T_m - T_n$, where the system of numbers T_i, named ***spectral terms***, defines the atomic system considered. The significance of this principle was understood by ***Niels Henrik David Bohr*** (1885 – 1962), who identified the numerical system T_i with the system of possible energy levels E_i of atomic electrons, with Planck's h constant as proportionality constant, namely $h\nu = E_n - E_m$, based on which he lay the foundation of the study of the quantum structure of atoms and of their interaction with light
1913 (Phil. Mag. 26, 1, 476, 857 (1913); ***The Theory of Spectra and Atomic Constitution***, Cambridge, 1922; ***On the Application of the Quantum Theory to Atomic Structure***, Cambridge, 1924). ***Albert Einstein*** (1879 – 1955) penetrated even deeper into the subtle mechanism of the interaction between light and atomic systems, in his fundamental work regarding the process of ***stimulated emission***, which, together with the processes of spontaneous emission and absorption allowed the

surprisingly simple and general deduction of the law of distribution of
radiant energy from the spectrum of dark bodies (***Zur Quantentheorie***
1917 ***der Strahlung***, Physicalische Zeitschrift, 18, 121, 1917). Whole decades
would pass before Einstein's ideas would be put to use in the quantum
apparatuses so widespread today, that amplify light through the
stimulated emission of radiation. As is known, the first laser ray sprung
1960 from a ruby crystal as late as June of 1960, in the laboratory headed by
Theodore H. Maiman (born 1927) at Hughes Aircraft Co. in Malibu,
California, thus bringing on the dawn of a new era in the millennial
history of optics (for a detailed historical presentation, see Mario
Bertolotti, ***Masers and Lasers - An Historical Approach***, Adam Hilger
Ltd. Bristol, 1983). In fact, a whole volume could be dedicated to the
quantum optics revolution triggered in 1900 by ***Max Karl Ernst Ludwig***
Planck (1858 – 1947), upon his introduction of the concept of ***light***
quantum and his establishing the law of thermal radiation (***Über***
1900 ***irreversible Strahlungsvorgänge***, Ann. Physik, 1, 69, 1900).

Let us go back to the moment in the history of geometrical optics
marked by Irish mathematician, astronomer, and physicist ***Sir William***
Rowan Hamilton (1805 – 1865), with his series of works entitled ***Theory***
of Systems of Rays, published in Transactions of the Royal Irish
1828 Academy (15, 69-174, 1828; 16, 1-61, 1830; 16, 93-125, 1831; 17, 1-144,
1837), in which he introduced his famous ***characteristic functions V, W,***
T of optical systems, based on Fermat's principle and variational
calculus. Thus, for example, the six variable function V, known today as
the ***point eikonal***, is defined as the optical path

$$V(x,y,z;x',y',z') = \int_{P}^{P'} n ds,$$

between the point $P(x,y,z)$ in object space and $P'(x',y',z')$ in image
space, which satisfies the eikonal equations

$$\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2 = n^2(x,y,z),$$

$$\left(\frac{\partial V}{\partial x\prime}\right)^2 + \left(\frac{\partial V}{\partial y\prime}\right)^2 + \left(\frac{\partial V}{\partial z\prime}\right)^2 = n'^2(x',y',z'),$$



where the partial derivatives represent the direction of the light beam in the point considered. Since the properties of optical systems can be described relative to the points and/or rays (direction cosines) in object and image spaces, Hamilton also introduced the "auxiliary" characteristic functions ***W*** and ***T***, which modern terminology calls the ***mixed eikonal*** and the ***angular eikonal***, respectively. Any of Hamilton's characteristic functions describe the optical system, and their usage offers specific advantages in various applications, as he himself demonstrated in the case of lens, mirrors, and revolving systems in general, in propagation through anisotropic media and the atmosphere.

The formalism devised by Hamilton for the light beams of geometrical optics was extended to the study of particle trajectory in classical mechanics, in a short work entitled ***On the Application to Dynamics of a General Mathematical Method Previously Applied to***
1834 ***Optics***, published in the British Association Report (1834), and in the definitive article ***On a General Method in Dynamics: by which the Study of the Motions of All Free Systems of Attracting or Repelling Points is Reduced to the Search and Differentiation of One Central Relation or Characteristic Function***, this time published in the most prestigious and well circulating journal (Phil. Trans. of the Royal Society, 1834).

The general and quite fertile mathematical method of characteristic functions, introduced by Hamilton in geometrical optics and in mechanics, constituted one of the most profound discoveries of the 19th century. But, whereas the Hamiltonian formalism in mechanics quickly became well known, thanks to the works of ***Karl Gustav Jacob Jacobi*** (1804-1851), Hamilton's great writings on geometrical optics was forgotten for decades (except by Maxwell (***On the Application of Hamilton's Characteristic Function***), Proc. London Math. Soc., 6, 117, 1875) and Thiesen (Ann. d. Physik, 45, 821, 1892)). The concept of characteristic function known as the ***eikonal*** was rediscovered and circulated once again by ***H. Bruns*** (***Das Eikonal***, K. sächs. Ges. d. wiss.
1895 Abhand. math. - phys. Kl., 21, 323-436, 1895), who wrote his work, extraordinarily, completely unaware of the existence of the work on optics written by his great precursor, as can be deduced from the sentence of page 329: "Eine ganz ähnliche Rolle, wie der Hamilton'sache Ansatz in der Mechanik, spielt nun der Eikonalbegriff auf dem allerdings weit engeren Gebiete der geometrischen Optik". Here we



have an amusing example of how science sometimes progresses in the dark. In fact, Bruns followed the path backwards, that is, from Hamilton and Jacobi's mechanics to geometrical optics, but starting from Malus's theorem (1808), and not from Fermat's more general principle (also applicable in anisotropic media). At any rate, Bruns obtained eikonal functions that were apparently simpler (having only four variables), but which actually described only particular instances of the Hamiltonian characteristic functions, with the light beam congruencies intersecting a pair of reference planes ($z = 0$, $z' = 0$). For this reason, a hundred years after the creation of Hamiltonian formalism in optics (1824 – 1844), the relative importance of Hamilton's and Bruns's contributions was still the object of a fierce debate between John Lighton Synge (***Hamilton's Method in Geometrical Optics***, J. Opt. Soc. Amer., 27, 75, 1937; ***Hamilton's Characteristic Function and Bruns'Eikonal***, J. Opt. Soc. Amer., 27, 138, 1937) and ***Maximillian Jakob Herzberger*** (***On the Characteristic Function of Hamilton, the Eikonal of Bruns and Their Use in Optics***, J. Opt. Soc. Amer., 26, 177, 1936; ***Hamilton's Characteristic Function and Bruns Eikonal***, J. Opt. Soc. Amer., 27, 133, 1937), which ended in favor of the former, but with the shorter term of eikonal retained. At the same time, Hamilton's compatriots, A. W. Conway and J. L. Synge, edited two volumes of his published work and manuscripts, under the titles ***The Mathematical Papers of Sir William Rowan Hamilton, vol. I: Geometrical Optics*** (1931) and Vol. II: ***Dynamics*** (1940), published by Cambridge University Press. In fact, the task of implementing and elaborating the ideas of ***Hamiltonian optics*** belonged almost exclusively to the 20th century, through such works as those as ***T. Smith*** (Trans. Opt. Soc., London, 23, 1921-1933), ***G. C. Steward*** (The Symmetrical Optical System, Cambridge, 1928), ***J. L. Synge*** (***Geometrical Optics, An Introduction to Hamilton's Method***, Cambridge, 1937, 1962), ***R. K. Luneburg*** (***Mathematical Theory of Optics***, Berkeley, 1964), ***M. J. Herzberger*** (***Modern Geometrical Optics***, Interscience, 1968), ***H. A. Buchdahl*** (***An Introduction to Hamiltonian Optics***, Cambridge, 1970), ***O. N. Stavroudis*** (***The Optics of Rays, Wavefronts and Caustics*** , Academic Press, 1972), ***T. Sekiguchi*** and ***K. B. Wolf*** (***The Hamiltonian Formulation of Optics***, Am. J. Phys., 55, 830, 1987), among others.

Generally, any system whose evolution is governed by

Hamilton's equations (see section 1.2, equation (64)) has numerous remarkable properties, such as ***Liouville's theorem***, according to which the volume elements in phase space are conserved. Let us consider the simple example of the one-dimensional optical problem, with beam trajectory $x(z)$ and direction of propagation $p_x(z)$, so that the volume element in phase space becomes the area element $dx.dp_x$. This

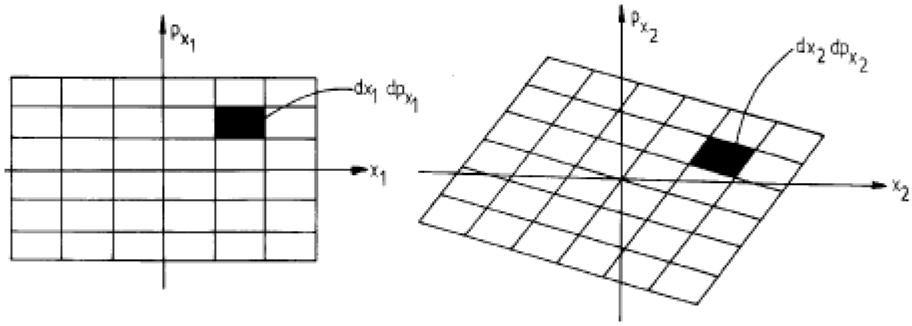


Fig. A.5 Conservation of the aria element (Liouville's theorem)

elementary area represents a narrow set of light beams that travel between points x and $x + dx$, with a direction of propagation between p_x and $p_x + dp_x$. According to Liouville's theorem, this area is conserved along the trajectory, so that, by considering any two points $P_1(z_1)$ and $P_2(z_2)$, we get $dx_1.dp_{x_1} = dx_2.dp_{x_2}$, as can be seen in Fig. A.5. Using an optical system, we can, for example, obtain a smaller breadth dx_2, but with a larger spread of beam directions dp_{x_2}, and conversely. This fundamental property of the phase space is a direct consequence of Liouville's theorem, and constitutes the essence of uncertainty relations in wave optics and quantum mechanics. Hamiltonian formalism paved the way for a general analogy between optics and mechanics, and served as a powerful tool for a dual description of natural phenomena, through trajectories and associated waves. As we know, the relevance of his work today is linked to the formulation of quantum mechanics theory and the representation of motion at microscopic levels.

The last decades have been marked by a series of new theoretical results based on which the classical field of geometrical optics has been considerably extended. Thus, ***J. Keller*** and his associates (Appl. Phys. Letters, 28, 426, 1957; J. Opt. Soc. Amer., 52, 2, 1962) have developed a geometrical theory of diffraction, starting from a generalization of Descartes' invariant. The results obtained constitute a marked progress from Kirchhoff's theory (1883), whose applicability is limited to



distances larger than the aperture area through which light diffracts (≫ λ). Starting off from a different direction, ***K. Miyamato*** and ***E. Wolf*** (J. Opt. Soc. Amer., 52, 615, 626, 1962) have elaborated on the ideas of ***B. B. Baker*** and ***E. T. Copson*** (***The Mathematical Theory of Huygens' Principle***, Oxford, 1939), and ***A. Rubinowicz*** (***Die Beugungswelle in der Kirchhoffschen Theorie der Beugung***, Warsaw, 1957), and have reached the conclusion that Kirchhoff's integral on the aperture area can be reformulated as a line integral on its outline. Thus, the initial explanation given by Thomas Young of the diffraction shapes is restored (he believed they represent the interference between the primary ray of incidence and the secondary waves generated by the aperture edges. ***M. Kline*** and ***I. W. Kay (Electromagnetic Theory and Geometrical Optics***, Interscience, 1965) have carried on ***R. K. Luneburg***'s work (***Mathematical Theory of Optics***, Berkeley, 1964), and have elaborated approximate methods that link the electromagnetic theory of light to geometrical optics and diffraction theory. Throughout the current volume, for the sake of simplicity, we have chosen to begin from the scalar equation of harmonic waves, for the deduction of the eikonal equation, as demonstrated by ***Arnold Sommerfeld*** and ***Iris Runge*** (***Anwendung der Vektor-rechnung auf die Grundlagen der Geometrischen Optik***, Ann. d. Physik, 35, 277,
1911 1911), as well as for the development of ***Gustav Robert Kirchhoff***'s scalar theory of diffraction (***Zur Theorie der Lichtstrahlen***, Ann. d. Physik, 18, 663, 1883). More recently, ***P. Hillion*** (J. Optics, 10, 11, 1979), has developed a theory which allows the description of polarized optical fields, by linearizing the eikonal equation. Another path of enquiry has been explored by ***D. Gloge*** and ***D. Marcuse*** (J. Opt. Soc. Amer., 59, 1629, 1969), who have started off from Fermat's principle and have achieved a quantification of light rays, by demonstrating that Gaussian beams are linked to wave packets of minimal spread. These ideas have been amply elaborated through the application of Lie groups and algebras in optics (***Lie Methods in Optics***, editors J. Sánchez-Mondragón and K. B. Wolf, in Lecture Notes in Physics, vol. 250, Springer, 1986), methods which have been also successfully used in quantum mechanics. Hamiltonian formalism thus proves capable of describing geometrical, wave-related and quantum properties, having deep roots in the simplectic geometry of phase space (***J. Sniatycki, Geometric Quantization and Quantum Mechanics, Springer***, 1980; ***V. Guillemin*** and ***S. Sternberg***, ***Symplectic***

Techniques in Physics, Cambridge, 1984). Finally, we note the recent successes in applying Lie groups to the study of higher order geometric aberrations.

While Hamilton was reviving the foundation of theoretical optics itself, and modelling geometrical optics and classical mechanics within the framework that one hundred years later would be taken up by quantum mechanics, considerable progress was being made in the field of instrumental and experimental optics. So it was that ***L. Schleiermacher*** initiated the theory of vignetting and proposed the method of least squares in optimizing the parameters of optical systems (***Über den Gebrauch der analytischen Optik***, Poggendorff Annalen, 14, 1828; ***Analytische Optik***, Darmstadt, 1842). The great mathematicians Cauchy and Gauss continued Euler's work in the field of optics, who, as we have shown, strived for many years to compensate color dispersion through combinations of lenses. Thus, Augustin Louis Cauchy (1789-1857) succeeded in giving a first good approximation of formula for dispersion in transparent media, in ***Mémoire sur la dispersion de la lumière***, 1836 (also refer to section 2.7, equation (262)), and introduced complex refraction indices to explain metallic reflection. ***Karl Friedrich Gauss*** (1777-1855) undertook systematic research regarding the properties of centered optical systems of paraxial rays, and introduced the concepts of conjugate and principal planes, which considerably facilitated the
1843 study of complex instruments (***Dioptrische Unterschungen***, 1843). ***Giovani Battista Amici*** (1786-1863) used the aplanatic property of Wierstrass points in the construction of objectives for microscopes of high numerical aperture (Ann. de chim. et phys., 12, 117, 1844), and introduced the method of immersion (also see section 2.1, Fig. 24). ***James Clark Maxwell*** (1831-1879) contributed to geometrical optics with the classical example of a perfect optical instrument with fisheye distribution – see section 3.3 (Cambridge and Dublin Mathematical
1854 Journal, 8, 188, 1854; Quart. Journ. of Pure and Applied Mathematics, 2, 233, 1858; for interesting generalizations, including Luneburg lenses, see R. Stettler, ***Optik***, 12, 529, 1955). Maxwell's profound work is only comparable to that of Newton and Einstein. As we know, it is to the genius of this great Scottish physicist that we owe electromagnetic theory and the equations that bear his name (***A Dynamical Theory of the***
1865 ***Electromagnetic Field***, Phil. Trans. Roy. Soc. London, 155, 459, 1865; ***A***

Treatise on Electricity and Magnetism, Oxford, 1873), the first great unification of electrical, magnetic, and optical phenomena, one of the greatest intellectual triumphs of all time. Let us here give an interpretation of the famous equation

$$c = 1/\sqrt{\varepsilon_0 \mu_0}$$

in Maxwell's own words: "The velocity of the transverse undulations in our hypothetical medium, calculated from the electromagnetic experiments of M. M. Kohlrausch and Weber (n.n. 1856), agree so exactly with the velocity of light calculated from the optical experiments of M. Fizeau (n.n. 1849), that we can scarcely avoid the inference that light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena", and further, " ... we have strong reason to conclude that light itself is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws". Soon
1888 afterwards (1888), Maxwell's work was crowned by the discovery of radio waves (which propagate at the same speed, and is reflected, refracted, interferes, diffracts, and is polarized just like ordinary light) by ***Heinrich Rudolf Hertz*** (1857-1894), who thus confirmed that light and electromagnetic waves share the same nature.

Whereas Fraunhofer opened the way for the precise correction of chromatic aberrations and the fashioning of modern refraction telescope objectives as early as 1817, solving the problem of geometric aberrations would await the arrival on the scene of Seidel and Petzval in the years
1843 1856-1857. Meanwhile, as we have seen, Gauss (1843) gave an elegant shape to the first order theory, in which Descartes' invariant $n\times \sin\theta$ is simply written as $n\times\theta$, a notation which implies a narrow bunch of light rays around the optical axis ($\theta \leq 0.1\, radians \cong 6°$). Under the spur brought on by the invention and perfection of photographic apparatuses and techniques (Niépce, 1826; Talbot, 1835; Daguerre, 1839; E. Becquerel and Draper, 1842; Foucault and Fizeau, 1845; Bond, 1850; De la Rue, 1860; Cros and Ducos du Haro, 1868; Eastmann, 1888; Lippmann, 1893; and many others) there emerged the new task of designing large aperture and field of view objectives, and, consequently, of extending optical system theory outside the paraxial domain. Obviously, we would get a better approximation if we retain the first two terms of the expansion $\sin\theta = \theta - (1/3\,!)\theta^3 + (1/5\,!)\theta^5 - (1/7\,!)\theta^7 + \cdots$, that is, if

we work in the framework of a third order theory. The deviation from the first order theory that occurs in this case will lead (see section 2.8) to the five primary aberrations (spherical, coma, astigmatism, field curvature, and distortion), also named Seidel aberrations, after ***Ludwig von Seidel*** (1821-1896), who first studied them in a systematic fashion (***Zur Dioptrik, über die Entwicklung der Gliedern 3-ter Ordnung***, Astron.
1856 Nachr., 43, 289, 305, 321, 1856). Seidel's analysis was subsequently simplified by various authors (for example, in the present work we have favored the method given in the unpublished notes of Edward L. O'Neill) and was extended, through various techniques, to the study of higher order geometric aberrations. As we have seen, the simple converging lens had long been used in the camera obscura (see della Porta, 1589), as well as in taking the first photograph (Niépce, 1826), but the first daguerreotypes were already being taken using achromatic doublets (Chevalier, 1830). But very long exposure times were used with small aperture diaphragms, to minimize geometric aberrations. Unlike regular telescopes, field curvature and distortion especially were no longer tolerated in the case of a good photographic objective. The first great theoretical and practical success was obtained by ***Josef Max Petzval*** (1807-1891), who studied the field curvature aberration in detail, deduced the requirements for image flattening (see section 2.8, equation (320)), and, based on previous calculations, created the fast photographic objective for portraits (***Bericht über die Ergebnisse einiger dioptrischer Untersuchungen***, Pesth, 1843; ***Bericht über optische Untersuchungen***, Ber. Keis. Akad. Wien, Math. - Naturwiss. Kl. 24, 50, 92, 129, (1857)). Petzval's ***"flattened" objective***, composed of two separate doublets, enabling high luminosity, but a small field of view, was perfected by Steinheil (1860) and Dallmeyer (1866), who, by making the separate doublets symmetrical, managed to also eliminate the distortion aberration, greatly enlarging the field of view at the same time.

Another remarkable achievement of this period was the discovery of the ***"schlieren"*** or ***"knife-edge"*** method (see Chapter III) by ***L. Foucault*** (***Mémoire sur la construction des télescopes en verre argent
1859 é***, Ann. de l'Observatoire Imp. de Paris, 5, 197, 1859) and, independently, by ***A. Töpler*** (***Beobachtungen nach einer neuen optischen Methode***, Bonn, 1864; Pogg. Ann. Physik u. Chem., 127, 556, 1866).

Before moving on to the "Abbe moment", let us note several more

advancements in the history of 19th century optics. Renowned astronomer ***Sir George Biddell Airy*** (1801-1892) was probably the first to correct astigmatism, by using a sphero-cylindrical lens to correct his own myopic astigmatism (1825). Widespread application of this method in ophthalmological practice would occur after 1862, the year in which ***Franciscus Cornelius Donders*** (1818-1889) published his treatise regarding cylindrical lens and astigmatism. Airy studied the formation of images through telescopes and the precision of astronomical observations in depth. Among his more than 500 published works, of special importance for determining the limit of applicability of geometrical optics and the resolution capabilities of optical instruments was the paper in which he calculated Fraunhofer diffraction through a circular aperture (***On the Diffraction of an Object Glass with Circular*
1835 ***Aperture***, Trans. Cambridge Phil. Soc., 5, 283, 1835). The maximal center, in which 83.9% of diffracted energy is concentrated, is a bright circular spot (***the Airy disk***) of angular radius γ, given by the famous formula

$$\sin\gamma = 1.22\,\frac{\lambda}{D}.$$

To him are also attributed the wave theory of the rainbow (see ***F. Uliu, Istoria curcubeului - De la Noe la Mie***, EMIA and UNIVERSITARIA Press, Deva-Craiova, 2005), as well as the Airy function in multiple beam interference theory.

Lord Rayleigh, John William Strutt (1842-1919) was the first to introduce a practical and simple criterion ($\gamma_{min} = 1.22\,\lambda/D$, see formula (249)) for describing the resolving power of optical instruments for incoherent light sources (Investigations in Optics with special reference
1879 to Spectroscope, Phil. Mag., 8, 261, 1879). Finally, we note the two simple criteria of stigmatism (see Chapter I, section 1.3, equation (98) and equations (102) and (112), respectively), namely ***the axial stigmatism condition***, deduced by ***Sir William Herschel*** (Phil. Trans. Roy. Soc., 111, 226, 1821) and the ***transversal stigmatism condition***, deduced first in relation to thermodynamics by one of the founders of that field, ***Rudolf Julius Emanuel Clausius*** (Pogg. Ann., 121, 1, 1864) and subsequently by renowned physiologist and physicist ***Hermann Ludwig Ferdinand von Helmholtz*** (Pogg. Ann. Jubelband, 557, 1874). However, the importance

of the latter condition in designing optical systems was signaled after its rediscovery by ***Ernst Karl Abbe*** (Jenaische Ges. Med. u. Naturwiss., 129, 1879).

The last third of the 19th century was dominated by the figure of optical physicist ***Ernst Karl Abbe*** (1840-1905), professor at the University of Jena, who, in close collaboration with microscope builder ***Carl Zeiss*** (1816-1888) and chemist ***Otto Schott*** (1851-1935), specialized in optical glass, lay the theoretical, technical and technological foundations of modern optical microscopy (see Ernst Abbe: ***Beiträge zur Theorie des Mikroskops under mikroskopischen Wahrnehmung***, Arch. f. mikr. Anat., 9, 413, 1873; ***Die optischen Hilfsmittel der Mikroskopie***, Braunschweig, 1878; ***Gesammelte Abhandlungen***, Gustav Fisher, Jena, 3 volumes, 1904 – 1906. A general exposition of Abbe's theory of optical instruments was published by his collaborator, S. Czapski, ***Theorie der optischen Instrumente***, 1893; also see S. Czapski, O. Eppenstein, ***Grundzüge der Theorie der Optischen Instrumente nach Abbe***, J. A. Barth, Leipzig, 1924). To illustrate the impact of these achievements, let us remember that microbiological and bacteriological research would have been impossible without microscopes with a resolving power close to the theoretical limit of diffraction. Like Fraunhofer in his time, Abbe was a unique combination of scientific genius, designer and inventor, who thus brilliantly demonstrated the fertile interaction between pure and applied science. However, unlike telescope and photographic objectives, production of performance microscopes had a late start, because of difficulties encountered in polishing such small lenses within the margins of acceptability, in significantly improving optical glass for achromatization, and in understanding the diffractive phenomena inherent to the observation of microbodies using correctly calculated systems involving such lenses. The history of the Abbe-Zeiss-Schott moment may be summarized with reference to the following events:

1846, Carl Zeiss is mechanic and lecturer at the University of Jena, and, following the suggestion of biologist J. Schleiden (1804 – 1881), begins building microscopes.

1861, Carl Zeiss receives the gold medal at the Industrial Exposition of Thuringia, for his high-quality microscopes, built, however, after the traditionally empirical fashion of trial and error.

1866 ***1866***, Carl Zeiss and the twenty employees of his small workshop in Jena

had produced around one thousand such microscopes, in the context of fierce competition from the company Hartnack in Paris, where water immersion objectives had been built since 1859. Zeiss begins collaboration with Abbe, who was then lecturer at the University of Jena, with the intent of building microscopes on a scientific basis. Initially, Abbe concentrated his efforts on introducing numerous measuring and control instruments in Zeiss's shop, particularly:

1867, Abbe's focometer, for measuring the focal distances of the objective and component lenses;

1869, Abbe's refractometer, published in Jenaische Ges. Med. u. Naturwiss, 8, 96, 1874, for determining the refractive indices of glass and liquid samples following measurements of the total reflection critical angle; the same year saw the introduction of the Abbe condenser, for illuminating samples with any angular aperture γ_1 in the maximal possible interval (±90°);

1870, Abbe's light meter, for determining the ***numerical aperture*** $AN = n_1 \sin \gamma_1$ of microscope objectives, where $n_1$is the refractive index of the immersion medium (air, water, oil, etc.) placed between the object and the frontal lens of the objective, and γ_1 is the angle between the marginal ray and the optical axis. The concept of numerical aperture was introduced by Abbe because, after numerous experiments, he found that this factor controls the brightness of the image and the resolving power of microscope objectives (close conjugate planes), unlike the ***number*** $= f/D$, which is the relevant factor when objects are far apart (telescope, photographic objective). Abbe demonstrated, first experimentally, then based on diffraction theory, that the minimal distance $(\delta r_1)_{min}$ between two points on the object that can still be resolved in the image (that is, the inverse of the spatial resolution) is proportional to λ_0, and inversely proportional to AN (see section 2.1, equation (123)). The immediate consequence of this research would be that, beginning with 1872, the Zeiss company would become famous for the performance of its microscopes, the first based on a correct theory and mathematical calculations. Abbe becomes Zeiss's partner (1875).

1871, Abbe publishes his studies concerning the intensity of light in optical instruments, in which he elaborates theoretical knowledge of diaphragms and the human pupil (Jenaische Ges. Med. u. Naturwiss., 6, 263, 1871). This research would later be further elaborated by ***M. von***

Rohr (Zentr. Ztg. Opt. u. Mech., 41, 145, 159, 171 (1920)).

1873 ***1873***, he publishes his fundamental work "***Contributions to the Theory of the Microscope and of Microscopic Perception***", in which he lays ***the theoretical foundations concerning diffraction and the formation of images***. Such a theory became necessary in the case of observing microbodies at a level of detail of the order of wavelengths, in which the effect of diffracted light can no longer be neglected. In order to illustrate Abbe's theory, let us consider such an object the size of a circular aperture P_1P_2, illuminated by a plane wave normally incident on the

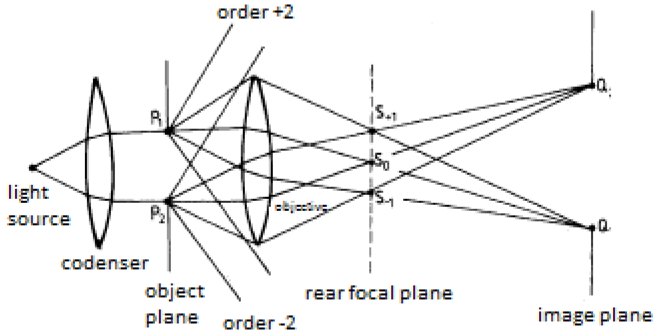


Fig. A.6 Abbe's theory of image formation in a telescope (circular aperture)

object plane (see Fig. A.6). The waves diffracted by the object are first focalized by the objective system on its rear focal plane, where it forms the corresponding Fraunhofer diffraction pattern, with the (spectral) maxima of various orders in S_0 , $S_{\pm1}$, $S_{\pm2}$, etc., and then further propagate, interfere, and finally form the inverted object image on the image plane of the objective. It becomes evident that the image obtained is proportionately clearer to the value of the angular aperture γ_1 (or, more generally, the numerical aperture $AN = n_1 \sin\gamma_1$) of the objective. The higher the angular aperture, the more spectra (spatial frequencies) contribute to the image's formation. This is how Abbe explained his famous formula for the theoretical limit of spatial resolution

$$(\delta r_1)_{min} = C\times \lambda_0/(n_1 \sin\gamma_1),$$

where the value of the constant C in the $0-1$ range ($C = 0.82$ for coherent illumination and $C = 0.61$ for incoherent illumination, see these details in ***M. Born, E. Wolf* , *Principles of Optics***, Pergamon, 1986, pp 418 – 424). As we have seen, the useful magnification of the



microscope is limited by the spatial resolution of the objective and by that of the observer's eye (see Chapter 1, section 2.5), with the ocular lens serving only to present to the eye under a convenient angle the (intermediary) image formed and resolved by the objective. Abbe confirmed his theory in numerous ingenious experiments (also see K. Michel, ***Die Grundlagen der Theorie des Mikroskops***, Stuttgart, 1950; K. Kranje, ***Simple Demonstration Experiments in the Abbe Theory of Image Formation***, Am. J. Phys., 30, 342, 1962), whose development led his successors to numerous important discoveries, such as the ***phase-contrast method***, developed by Dutch physicist ***Frits Zernike*** (1888 – 1966) in the years 1932 – 1934, for which he received the Novel prize for physics in 1953 (F. Zernike, ***Beugungstheorie der Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode***, Physica, 1, 689, 1934; Zs. Tech. Phys., 16, 454, 1935; Phys. Zs., 36, 848, 1935; Physica, 9, 686, 974, 1942; ***How I Discovered Phase Contrast***, Science, 121, 345, 1955; for a detailed presentation of the phase-contrast method, see M. Françon, ***Le contrast de phase en optique et en microscopie***, Paris, Revue d'Optique, 1950; A. H. Bennett, H. Jupnik, H. Osterberg, O. W. Richards, ***Phase Microscopy***, New York, Wiley, 1952). Zernike himself described his method as a logical application of Abbe's theory regarding the formation of images in the microscope in the case of transparent objects of irregular optical thickness (***phase objects***), such as those frequently encountered in biology and crystallography, that is, objects that modify the phase, but not the amplitude of the incident wave. The idea behind Zernike's method is to place a thin transparent plate (***phase plate***) across the objective's rear focal plane, so that the zeroth-order phase S_0 is advanced or retarded by $\pi/2$ in relation to the other diffraction orders, and, consequently, the invisible image of the (transparent) phase object becomes visible (contrasted), similar to that of an (absorbing) amplitude object; the phase differences of the phase object are thus transformed in the corresponding differences in brightness or intensity of the image. Since the objective's rear focal plane is usually located within the system of lenses, the phase plate is "incorporated" into the objective. The influence of Abbe's theory in our times is profound and fertile, his example of the microscope has allowed us to reach the concept of the ***spatial filter***, to understand the significance of the Fourier

transformations in image formation, and to develop a new field within optics, Fourier optics and the optical processing of information.

1879, Abbe establishes the general form of the necessary condition of aplanatism, or the ***sine condition*** that bears his name, later supplemented by ***F. Staeble*** (Münchener Sitz. - Ber., 183, 1919) and E. Lihotzky (Wiener Sitz. - Ber., 128, 85, 1919) with the isoplanatism condition. These conditions are major criteria for correction in designing modern optical systems.

1879 ***1879***, Abbe introduces the achromatic objective (for two wavelengths) with homogenous immersion (oil), with a numerical aperture $NA = 1.25$. However, further perfection of these objectives required the production of new types of glass that would allow combinations of glass with low refractive indices and high dispersion and glass with high refractive indices and low dispersion. Fortunately, in the year 1879, Abbe finds an ideal partner for solving the problem of new glasses in chemist ***Otto Schott***, together with whom, after several demanding years, he would build his famous ***apochromatic objectives***, which satisfy achromatization conditions for three colors and the sine condition
1886 (aplanatism) for two colors (Jenaische Ges. Med. u. Naturwiss., 1886). Abbe's apochromatic objective (an immersion objective, composed of ten lenses, $f = 2mm$, $NA = 1.4$), supplied by the Zeiss company beginning with 1886, would inaugurate a new era for the finest of visual observations, and for microphotography, its resolving power reaching the theoretical diffractive limit.

1884, the glass manufacturing company "Jenaer Glaswerke Schott und Genossen" is established, which in 1886 is already producing 44 types of optical glass, of unprecedented quality and variety.

1889 ***1889***, Abbe presents this highest performance achromatic objective, with immersion in monobromonaphtalene ($AN = 1.6$).

1889, after the passing of Carl Zeiss (1888), Abbe becomes sole owner of the firm and establishes the Carl Zeiss Foundation for scientific research and social betterment (Carl Zeiss Stiftung), thus becoming a great pioneer in the sphere of social reform.

Beside these apparatuses, there are many more devices and precision instruments linked to Abbe's name, such as the Abbe spectrometer (based on the principle of autocollimation), for quickly determining the refractive indices and dispersion of glass (1874), the



Abbe interferometer, for easily testing plane-parallel plates (1885), projecting oculars for microphotography (1886), an illumination system with mirror and lenses, known as the Abbe condenser (1886), the comparison microscope (1891), an improved version of Fizeau's interferential dilatometer, for determining the thermal dilation coefficient of glass (1893), the Abbe Porro prism system for image inversion in terrestrial telescopes (1895), the introduction within the Zeiss company of new departments for binocular telemetry, for photographic objectives and for astronomical telescopes. Next, Abbe further elaborated his ideas through his close collaborators, among whom let us mention ***Carl Pulfrich*** (1858 – 1927), the inventor of the stereocomparator, the stereo locator and the photometer that bears his name, and ***P. Rudolph***, who, by inventing "anastigmats" (1890), marked the birth of modern photographic objectives (such as the famous Tessar objective, devised by him in 1902). For a more detailed description of Abbe's numerous achievements in optics, see M. von Rohr Ernst Abbe, Fischer Verlag, Jena, 1940; N. Guenther, ***Ernst Abbe, Schöpfer der Zeiss - Stiftung***, Fischer Verlag, Stuttgart, 1951; F. Schomerus, ***Werden und Wessen der Carl Zeiss-Stiftung***, Fischer Verlag, Stuttgart, 1955; H. Volkmann, ***Ernst Abbe and his Work***, Appl. Optics, 5, 1720, 1966; ***Jenaer Rundschau*** (***Jena Review***), No. 3, 1973, an edition dedicated to the 100 year anniversary of Abbe's development of the theoretical and practical basis for modern optical microscopy.

Among the various methods of illuminating transparent objects, let us further mention A. Köhler's condenser (Zs. f. wiss. Mikroskopie, 10, 433, 1893; 16, 1, 1899), critical illumination (F. Zernike, Physica, 5, 794, 1938), and the recent use of aspherical condensers with very small numbers f which ensure a great density of luminous flux on the object (see the aspherical lenses presented in Chapter I, section 1.3).

The greatest contributions to the development of the fundamental geometrical optics of our century were made by ***Alvar Gullstrand*** (***Allgemeine Theorie der monochromatischen Aberrationen***, Acta Reg. Soc. Sci. Uppsala, 3, 1900; ***Die reele optische Abbildung***, Svenska Vetensk. Handl., 41, 1, 1906; ***Tatsachen und Fiktionen in der Lehre der optischen Abbildung***, Arch. Optik, 1, 1, 81, 1907; ***Das allgemeine optische Abbildungssystem***, Svenska Vetensk. Handl., 55, 1, 1915), ***Thomas Smith*** (***On Tracing Rays through an Optical System***, Proc. Phys. Soc. London,

28, 502, 1915; 30, 221, 1918; 32, 252, 1920; 33, 174, 1921; 57, 286, 1945), ***H. Boegehold*** (***Über die Entwicklung der Theorie der optischen Instrumente seit Abbe***, Ergeb. d. exakt. Naturwiss., 8, 1929; ***Raumsymmetrische Abbildung***, Z. Instrumentenk., 56, 98, 1936), ***M. Hertzberger*** (***Strahlenoptik***, Springer, Berlin, 1931; ***Modern Geometrical Optics***, Interscience, New York, 1968), and ***G. Slusarev*** (***Metodi di calcolo dei sistemi ottici***, Firenze, 1943; Geometriceskaia optika, Moskva, 1946). For his contributions in the field of ophthalmology (cornea astigmatism and abnormal shapes, correction lens for after the removal of the cataract crystalline), Gullstrand received the Nobel prize for physiology and
1911 medicine (1911). By introducing the method of matrix algebra in ray tracing and designing optical instruments, T. Smith became one of the great professors of the next generation of opticians. In his first work, Boegehold writes a synthesis of the most important achievements in geometrical optics of the first thirty years of the 20th century. Herzberger applies in geometrical optics Hamilton's fundamental ideas (also see his dispute with Synge), and develops mathematical models for optical systems. Slurasev undertakes an ample analysis of Seidel aberrations and the methods of optical calculations. Finally, in the fifth decade of the 20th century, ***B. R. A. Nijboer***, ***F. Zernike***, and ***N. G. van Kampen*** elaborated the theory of aberration diffraction (see M. Born and E. Wolf, Principles of Optics, Pergamon, 1986, Chapter IX).

As we have seen, the new optical glass introduced by Schott (for example, the crown glass with barium oxide (BaO) stabilizer, introduced into the composition as a nitrate ($Ba(NO_3)_2$) or barium carbonate ($BaCO_3$), which has a high refractive index and low dispersion), revolutionized the correction of dioptric systems displaying chromatic and geometric aberrations. So it is that today's photographic objectives are based on P. Rudolp's (the Zeiss company) and H.D. Taylor's (the Cooke company) ***anastigmats***, achromatized systems with corrected field curvature, astigmatism, and coma, for large angular fields, further perfected by W. Merté, R. Richter, M. von Rohr, ***Das photographische Objectiv***, Springer 1932; E. Wandersleb, ***Die Lichtverteilung im Grossen im der Brennebene des photographischen Objektivs***, Akademie Verlag, Berlin, 1952; J. Flügge, ***Das photographische Objektiv***, Springer, 1955.

Today, designing dioptric systems has become a matter of automation, in which matrix methods play an important role. Programs

that allow automated design can thus be written, for tasks ranging from simple ray tracing, to the design of systems of the highest performance, in which third, fourth, and even higher order aberrations are corrected. For an introduction, see D.P. Feder, ***Automatic Optical Design***, Applied Optics, 2, 1209, 1963.

The great majority of lens and mirror systems use spherical surfaces, which are easily produced with the required optical precision (tolerance $\ll \lambda$), but which pose the challenge of correcting inherent aberrations. However, there are also high performance optical instruments that contain elements of aspherical surfaces (Cartesian, toroidal, cylindrical), despite the difficulty with which they are manufactured (see T. Sakurai, K. Shishido, ***Study on the fabrication of aspherical surfaces***, Appl. Optics, 2, 1181, 1963). Usually, rigorous axial stigmatism of centered optical systems can be achieved with an single aspherical surface, and aplanatism, with two. Such a system (a telescope aplanatic lens with a large angular field, composed of two aspherical mirrors) was calculated by ***Karl Schwarzschild*** (***Theorie der Spiegeltelescope***, Abh. Königl. Gesellsch. d. Wiss. zu Göttingen, Math. - physik. Klasse,
1905 4, 1905), and was most often applied in microscopy (see D. S. Gray, ***A New series of Microscope Objectives***, Journ. Opt. Soc. Amer., 39, 723, 1949; R. C. Mellors, ***The Reflection Microscope***, Science, 112, 381, 1950). Although the principle of the reflection microscope (so, without chromatic aberrations) had been formulated by Newton and later taken up by Amici, since the objective of such a microscope is similar to the objective of the Newtonian telescope which functioned inversely, the idea was not materialized until ***Cecil Reginald Burch*** took up the task
1947 (Proc. Phys. Soc., 59, 41, 1947), starting from Schwarzschild's analytical solution for the two-mirror aplanatic system. Such a ***microscope reflection objective***, of high angular aperture, which reminds one of the Cassegrain telescope objective, is illustrated in Fig. A.7.

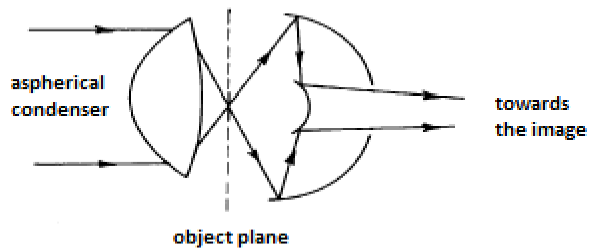


Fig. A.7 High angular aperture microscope reflecting objective

Once adjusted for the visual spectrum, this system also allows for microphotography within the ultraviolet spectrum, in which spatial resolving power is higher. Generally, catoptric systems are also widely used outside the field of optics, beginning with X-ray focalization (see V. P. Kirkpatrik, H. H. Pattee, Jr., ***X- Ray Microscopy***, in Encyclopedia of Physics, 30, 305 - 336, editor S. Flügge, Springer, 1957; H. Riesenberg, ***Über zentralabschattungsfreie, rotationssymmetrische Spiegel systeme mit besonderer Berücksichtigung ihrer Eignung als abbildende Optik für extrem weiche Röntgenstrahlen***, Jenaer Jahrbuch, II, 250 - 328, 1963), to observatory radio telescopes in Jodrell Bank (England) and Parkes (Australia). In the recent history of great reflection optical telescopes with parabolic primary mirrors, let us mention the installation at Mont Wilson, U.S.A., of a $D = 152\ cm$ telescope in 1908 (with a mirror polished by G.W. Ritchey) and a $D = 254\ cm$ telescope in 1918, the $D = 508\ cm$ telescope (designed by G.E. Hale, with the mirror polished and aluminized by J.D. Strong) installed in 1947 at the Mont Palomar observatory, U.S.A., and the recent installation of the largest reflection telescope, with a $D = 600\ cm$, at the Caucaz observatory. All these remarkable instruments illustrate the fertile interaction between science and technology. The development of the technology for aluminizing
1932 telescope mirrors through evaporation, by ***John Donovan Strong*** (1932), had a profound influence on astronomical observations. A significant progress was also achieved when catadrioptric telescopes with correcting refringent plate were built (see Chapter II, section 2.6, Fig. 61) by ***Bernhard Voldemar Schmidt*** (Central Zeitung f. Optik u. Mechanik,
1931 52, 1931; Mitt. Hamburg Sternwarte Bergedorf 36, 15, 1932; also see R. Schorr, Zs. f. Instrum., 56, 336, 1936; Astr. Nachr., 258, 45, 1936; Mitt. Hamburg Sterwarte Bergedorf, 42, 175, 1936; C. Carathéodory, Elementare Theorie des Spiegelteleskops von B. Schmidt, Teubner, Leipzig, 1940; E. H. Linfoot, Recent Advances in Optics, Chapter IV, Clarendon, Oxford, 1955) and ***Dmitri Dmitrievici Maksutov*** (***Novîie Katadioptriceskie Sistemî***, Dokl. Akad. Nauk S. S. S. R., 37 - 127, 1942; ***New Catadioptric Meniscus Systems***, Journ. Opt. Soc. Amer., 34, 270, 1944; also see A. Bouwers, ***Achievements in Optics***, Elsevier, New York, 1950). These high-performance instruments, of very large angular fields, in the Gregory or Cassegrain versions, allowed the astronomical mapping of the whole sky (see, for example, the so-called ***Sky Survey***



effected on the boreal sky with the use of the Schmidt $D = 180\ cm$ mirror and 120 cm diameter correcting plate telescope at the observatory of Mount Palomar). In the field of microscopy, let us recall D.S. Gray's perfecting the Burch objective, with his catadioptric version (op. cit., 1949); he significantly increased the angular aperture by adding a frontal system of molten quartz and fluorene lenses (transparent in the ultraviolet spectrum). More recently, the use of high-speed computers has allowed radical improvements in the design of complex optical systems with aspherical lenses for the most diverse of applications (remote sensing, remote control, tracing), with performances nearing the natural limit of diffraction. Extreme precision has been achieved in polishing optical elements by way of ionic bombardment. The application of simple and multiple coating (reflective, anti-reflective coating) has become a common process. The technology of infrared materials has reached great heights. Plastic materials have been introduced in the production of optical elements (prisms, lenses, aspherical lenses, grating replicates, optical fibers). Ceramic glasses have been discovered, with extremely low dilatation coefficients.

An older idea, patented by ***John Logie Baird*** (British patent 285738, 15th of February 1928), that predicted the possibility of transmitting light and images through transparent dielectric fibers, was
1969 taken up once more by ***Charles Kuen Kao*** (1966), so that, in 1969, the Corning Glass company was already producing the first glass ***optical fibers***, with relatively small losses ($\cong 20\ dB/km$), thus inaugurating the era of optical fiber communications (see N.S. Kapany, ***Fiber Optics, Principles and Applications***, Academic Press, New York, 1967; D. Gloge, ***Optical Fibers for Comunications***, Appl. Optics, 13, 249, 1974; D. Marcuse, ***Principles of Optical Fiber Measurements***, Academic Press, New York, 1981; A. B. Sharma, S. J. Halme, M. M. Butusov, ***Optical Fiber Systems and their Components***, Springer, Ser. Opt. Sci., 24, 1981; ***Y. Suematsu, K. Iga***, Introduction to Optical Fiber Communications, Wiley, New York, 1982; A. H. Cherin, ***Introduction to Optical Fibers***, McGraw-Hill, New York, 1983). The array of diameters of fibers used today for leading light across large distances spans between several to thousands of microns (a human hair has a diameter of around 50 microns). If the diameter of the fiber is large compared to the wavelengths, light propagation may be treated in geometrical optics terms, as we have done

in Chapter III, section 3.2, in the case of cylindrical structures. If, however, the diameter is comparable to the wavelength, light propagates through the fibers as through a waveguide of optical frequency ($\cong 10^{15}\ Hz$), in which case rigorous electromagnetic theory must be applied; the same rule applies when studying light propagation through very thin dielectric strata. This lead to the start of new chapter
1969 in applied optics, termed, in short, ***integrated optics*** (S.E. Miller, 1969). Like hollow metallic guides in the case of microwaves, rigorous analysis of the luminous electromagnetic wave propagation is carried out using Maxwell's equations and the corresponding limit conditions. Wavelength in the optical spectrum is around 10^4 smaller than in the microwave spectrum, and the advantages of optical frequencies and of the corresponding optical circuit and guide miniaturization are numerous (see M.J. Adams, ***An Introduction to optical Waveguides***, Wiley, New York, 1981; R.G. Hunsperger, ***Integrated Optics, Theory and Technology***, Springer, Ser. Opt. Sci., 33, 1984).

Here end our considerations regarding the history of geometrical optics, with the observation that, although this type of optics today seems to us a limiting case of Maxwell's equations applied to the phenomena of propagation of electromagnetic fields of very small wavelength, designing optical instruments is still based on tracing light rays through the systems considered, since it seldom happens that diffraction goes beyond geometric aberrations. Likewise, the study of finer phenomena, of a wavelike nature, such as interference, diffraction and polarization, always entail preliminary analysis of the geometrical path of light beams. This is because, in order to obtain a first good approximation of the propagation of light, we do not require any hypothesis regarding its "ultimate" nature. Purely geometrical representations suffice, a fact which Fermat, in fending off attacks from the Cartesians' side, summed up in the following words:

"... je ne prétends ni n'ai jamais prétendu être de la confidence secrète de la Nature. Elle a des voies obscures et cachées que je n'ai jamais entrepris de pénétrer; je lui avais seulement offert un petit secours de géometrie au sujet de la réfraction, si elle en ait eu besoin. "

("... I do not claim, nor have I ever claimed to be among Nature's confidents. Her ways are concealed and unexplained, and I have never attempted to decipher them; I have merely offered her some minor

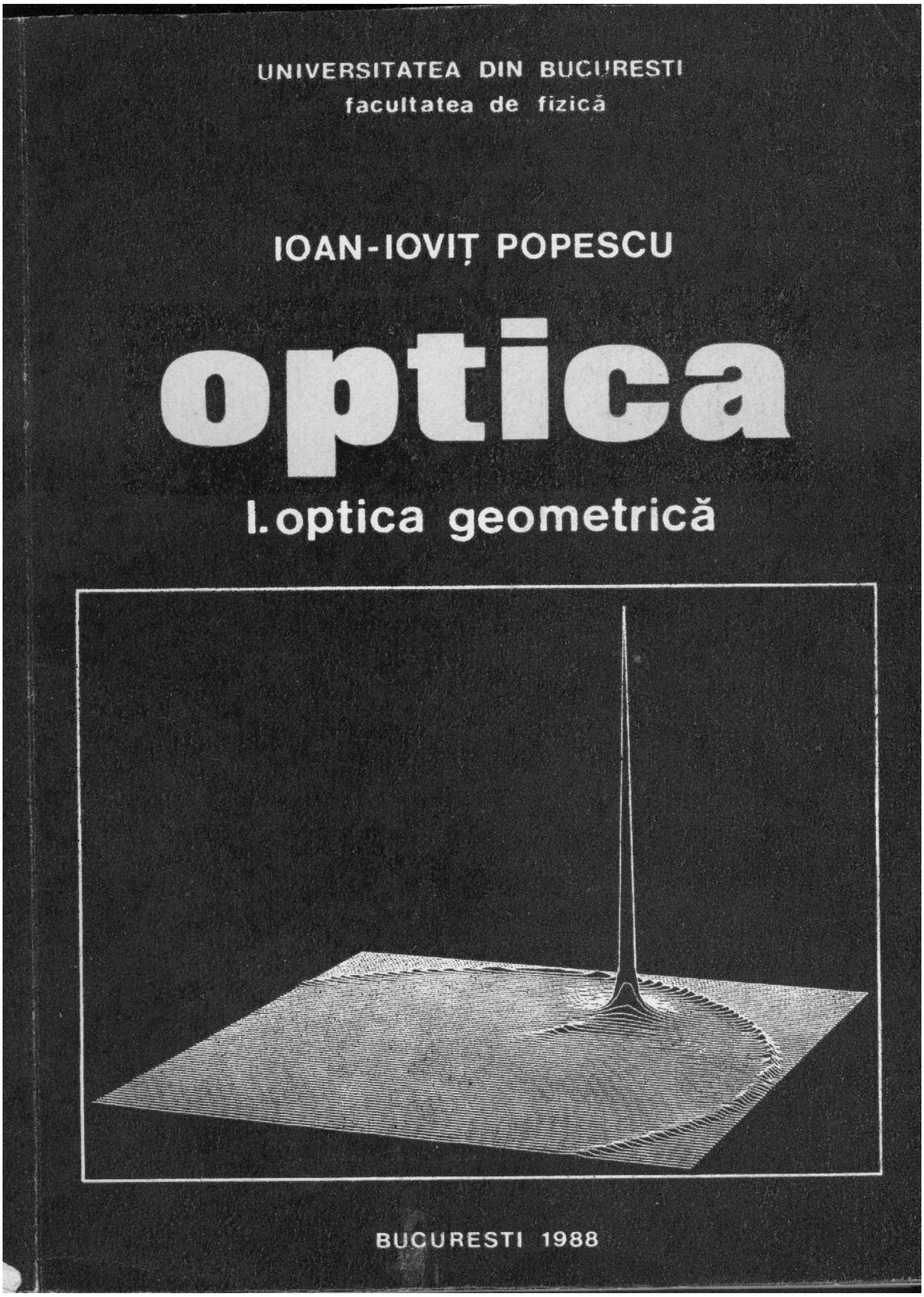


assistance in geometry, regarding refraction, that is, if she ever needed assistance as such.")

Let this minimal program, brilliantly achieved through Fermat's principle and Hamilton's work of geometrical optics, constitute a final ***motto*** for the present book.

Appendix B

SAMPLE OF ORIGINAL HANDWRITTEN TEXTBOOK



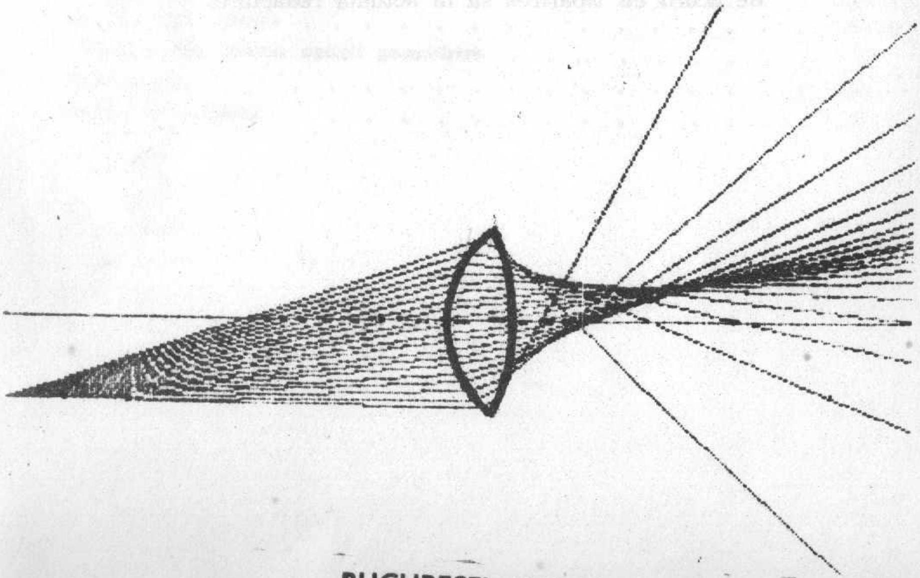


UNIVERSITATEA DIN BUCUREŞTI
facultatea de fizică

IOAN-IOVIŢ POPESCU

optica

I. optica geometrică



BUCUREŞTI 1988

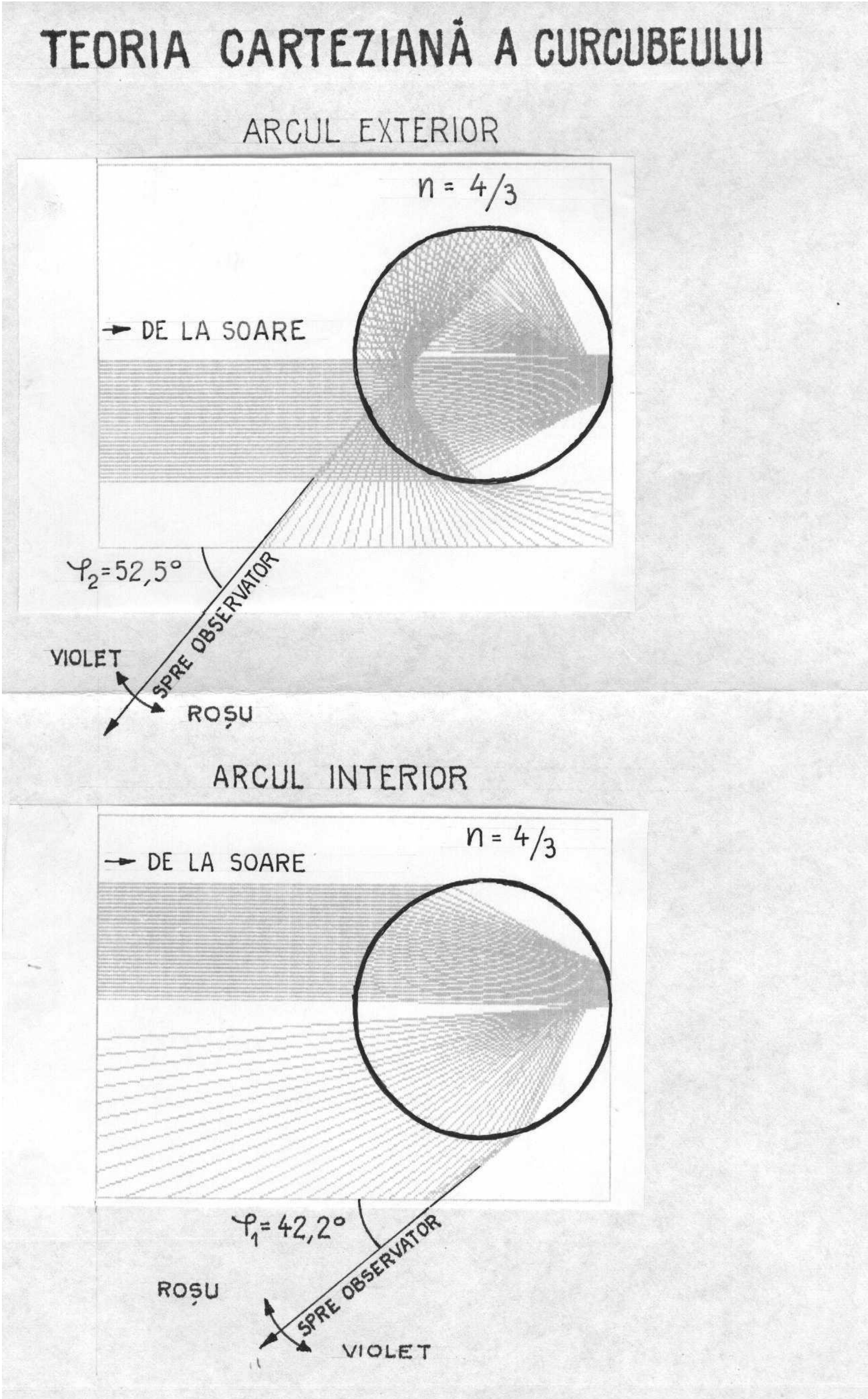
Contemporary Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017



OPTICA GEOMETRICĂ
IOAN-IOVITZ POPESCU
1988

CUPRINS

Contemporary Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017



Contemporary Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017

Îndice de subiecte

http://editura.mttlc.ro
The University of Bucharest. 2017

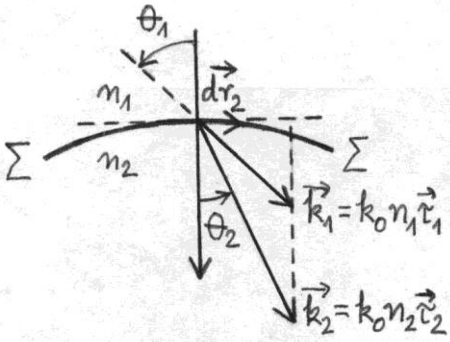


Fig. 7

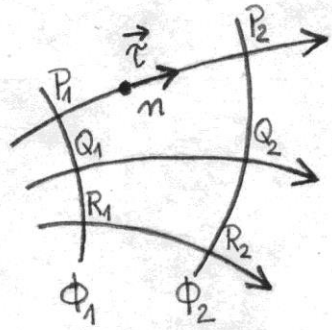


Fig. 8

Să aplicăm, în continuare, teorema invariantului integral al lui Lagrange, ec.(1.21), pentru cazul particular în care conturul de integrare este chiar traiectoria unei raze de lumină astfel că $n\vec{\tau}\cdot d\vec{r} = n\vec{\tau}\cdot\vec{\tau}\,ds = n\,ds$. În acest caz integrala (1.21) între două puncte P_1, P_2 oarecare ale razei, notată cu $[P_1P_2]$, poartă numele de *drum optic* și are următoarele expresii echivalente

$$(26) \qquad [P_1P_2] \stackrel{def}{=} \int_{P_1}^{P_2} n\,ds = \lambda_0 \int_{P_1}^{P_2} \frac{ds}{\lambda} = c(t_2-t_1) = \Phi_2 - \Phi_1 \quad ,$$

unde am folosit expresia $n = k/k_0 = \lambda_0/\lambda$ și relația $n\,ds = c\,dt$, ec.(1.12). Cu alte cuvinte, *drumul optic între două puncte ale unei raze de lumină este* [egal cu drumul $c(t_2-t_1)$ parcurs în vid și] *proporțional cu numărul de lungimi de undă, cu timpul de propagare a luminii, respectiv cu diferența de fază între oscilațiile armonice ale câmpului optic în punctele considerate.*

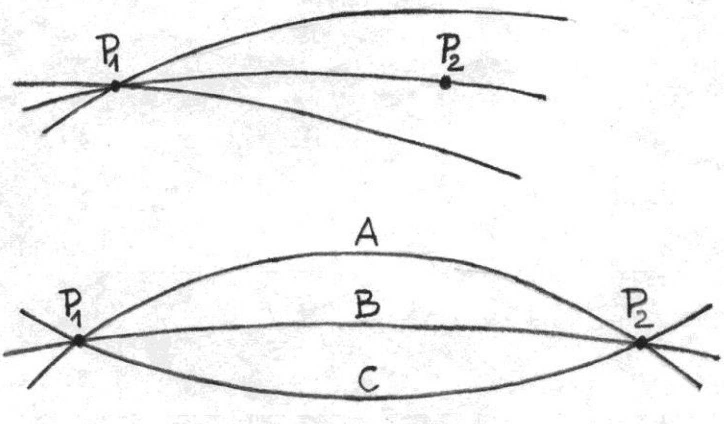
Conceptul de drum optic permite să formulăm următoarea proprietate generală, denumită *principiul egalității drumurilor optice* (teorema Malus-Dupin), conform căreia, indiferent de mediile optice și de suprafețele de discontinuitate străbătute, *drumul optic între două suprafețe de undă oarecare este același pentru toate razele de lumină*. Valabilitatea acestei aserțiuni rezultă

1.3 Condiții generale de stigmatism

Să considerăm un fascicul conic (homocentric) de raze de lumină emis de o sursă punctuală P_1 (Fig.1.14,a). În general, din infinitatea de raze a acestui fascicul, numai una singură va trece printr-un alt punct P_2, și anume traiectoria extremală care satisface principiul lui Fermat. Pe de altă parte, funcția ideală a instrumentelor optice de format imagini constă în dirijarea fasciculului de raze în așa fel încît fiecărui punct P_1 din spațiul obiectului să-i corespundă un singur punct P_2 în spațiul imaginii. Din acest motiv, în continuare ne vor interesa acele cazuri excepționale în care punctele P_1 și P_2 sînt legate printr-o infinitate de raze (Fig.1.14,b).

<u>Stigmatismul</u> reprezintă conceptul fundamental al teoriei geometrice a imaginilor optice. Denumirea provine din cuvîntul grecesc στιγμα care înseamnă punct. Prin definiție, un sistem optic este stigmatic sau punctual pentru perechea de puncte P_1, P_2 dacă un fascicul conic de raze cu vîrful în P_1 este transformat într-un fascicul conic de raze cu vîrful în P_2.

(a)

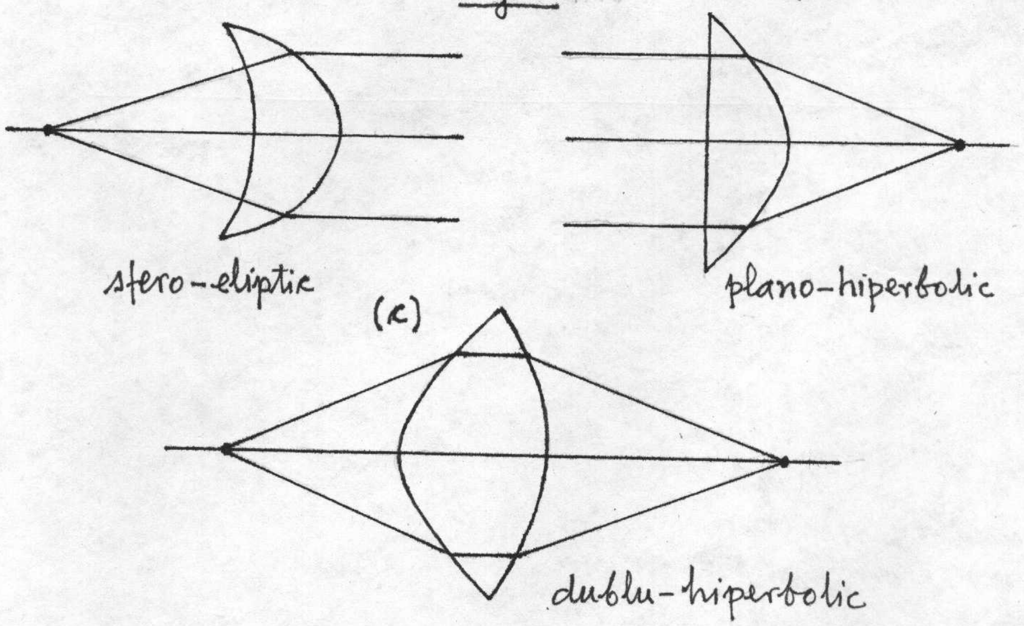


(b)

Fig.14

Contemporary Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017

Fig.18



Deasemenea, proprietatea de stigmatism riguros a suprafețelor carteziene de refracție este folosită pentru realizarea lentilelor asferice. În principiu, referindu-ne la Fig.1.15, a, b, c, o lentilă asferică, confecționată din mediul optic n_2, este limitată de suprafața carteziană Σ și orice suprafață sferică Φ_2 cu centrul în P_2. In practică sînt folosite suprafețele carteziene cu secțiune conică, cum este ilustrat în Fig.1.18 pentru lentila sfero-eliptică (a), plano-hiperbolică (b) sau dublu-hiperbolică (c). Datorită lipsei aberației sferice (v. § 2.8), lentilele asferice pot avea diametre D ale aperturii mult mai mari și distanțe focale f mult mai mici decît lentilele sferice. În consecință, se poate ajunge la numere $f \overset{def}{=} f/D$ (v. § 2.6) foarte mici (în practică pînă la 0,6), respectiv la o densitate de flux luminos în planul imaginii foarte mare. Lentilele asferice permit astfel folosirea cea mai eficientă a surselor și detectorilor de lumină, de unde și numeroasele lor aplicații în sistemele optice actuale de comunicații și control. În fine, mai remarcăm

Contemporary Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017



Bibliography

1. S. A. Ahmanov, S. Yu. Nikitin, ***Physical Optics***, Clarendon Press, Oxford, 1997;
2. J. A. Arnaud, ***Beam and Fiber Optics***, Academic Press, New York, 1976;
3. Bai Gui-ru, Guo Guang-can, ***Problems and Solutions on Optics***, World Scientific, Singapore, 1991;
4. V. V. Bianu, ***Optică geometrică***, Editura Tehnică, București, 1962;
5. J. W. Blaker, ***Geometric Optics. The Matrix Theory***, Marcel Dekker, Inc., New York, 1971;
6. M. Born, E. Wolf, ***Principles of Optics***, Pergamon Press, Oxford, 1986;
7. A. Boussié, ***Physique- exercices avec solutions***, Vuibert, Paris, 1984;
8. G. A. Boutry, ***Instrumental Optics***, Hilger, London, 1961;
9. C. Bozan, ***Curs de Optică*** (p. I), Tipografia Universității din Timișoara, 1975;
10. G. G. Brătescu, ***Optica***, Editura Didactică și Pedagogică, București, 1982;
11. G. Brooker, ***Modern Classical Optics***, Oxford University Press, New York, 2003;
12. W. Brouwer, ***Matrix Methods in Optical Instrument Design***, W. A. Benjamin, Inc. New York, 1964;
13. H. A. Buchdahl, ***Optical Aberration Coefficients***, Oxford University Press, London, 1954; ***An Introduction to Hamiltonian Optics***, Cambridge University Press, 1970;
14. E. J. Butikov, ***Optika***, Izd. Vîssaia școla, Moskva, 1986;
15. G. Chartier, ***Intoduction to Optics***, Springer, Berlin, 2005;
16. A. H. Cherin, ***Introduction to Optical Fibers***, Mc Graw - Hill, New York, 1983;
17. H. Chrétien, ***Calcul des combinaisons optiques***, Masson, Paris, 1959;
18. A. E. Conrady, ***Applied Optics and Optical Design***, Dover Publication, Inc., New York, 1975;
19. S. Cornbleet, ***Microwave Optics***, Academic Press, New York, 1976;
20. F. S. Crawford jr., ***Unde*** (Cursul de Fizică Berkeley, vol.III), Editura Didactică și Pedagogică, București, 1983;
21. René Descartes, ***La Dioptrique***, Oeuvres de Descartes, Vol. VI, C. Adam et P. Tannery (éditeurs), Paris, 1902;
22. F. Desvignes, ***Détection et détecteur de rayonements optiques***, Collection Mesures Physiques, Paris, 1987;
23. R. W. Ditchburn, ***Light***, Dover Publications Inc., New York, 1991;
24. P. Dodoc, ***Calculul și construcția aparatelor optice***, Editura Didactică și Pedagogică, București, 1983;
25. D. O. Dorohoi, ***Optica***, Editura Ștefan Procopiu, Iași, 1995;
26. E. Elbaz, F. Roux, ***Optique matricielle***, Ellipses, Paris, 1989;
27. G. R. Fowles, ***Introduction to Modern Optics,*** Dover Publications, New York, 1975;
28. A. Gerrard, I. M. Burch, ***Introduction to Matrix Methods in Optics***, John Wiley & Sons, New York, 1975;
29. A. Ghatak, K. Thyagarajan, ***Contemporary Optics***, Plenum Press, New Delhi, 1978;
30. R. Grigorovici, ***Curs de optică***, partea I, ***Optica Geometrică***, Universitatea București, 1957;
31. R. D. Guenther, ***Modern Optics***, Wiley, New York, 1990;



32. W. R. Hamilton, ***The Mathematical Papers of Sir William Rowan Hamilton***, Vol.I, ***Geometrical Optics***, edited by A. W. Conway and J. L. Synge, Cambridge University Press, New York, 1931;
33. O. S. Heavens, R. W. Ditchburn, ***Insight into Optics***, Wiley, New York, 1991;
34. E. Hecht, ***Optique, Cours et Problèmes***, McGraw-Hill, Inc., New York, 1980;
35. E. Hecht, A. Zajac, ***Optics***, Addison - Wesley Publ. Comp., Inc., Reading, Mass., 1987;
36. E. Hegedüs, ***Introducere în Optică*** (I, II), Tipografia Universității din Timișoara,1974;
37. G. Huțanu, ***Zigzag în lumea opticii***, Ed. Albatros, București, 1986;
38. M. Herzberger, ***Modern Geometrical Optics***, Interscience Publishers, New York, 1968;
39. Christian Huygens, ***Traité de la lumière***, Van der Aa, Leiden, 1906, republicat de Gauthier - Villars, Paris, 1992;
40. K. Iizuka, ***Engineering Optics***, Springer - Verlag, Berlin, 1985;
41. I. Iova, ***Elemente de optică aplicată***, Editura Științifică și Enciclopedică, București, 1977;
42. F. A. Jenkins, H. E. White, ***Fundamentals of Optics***, McGraw-Hill, Inc., 1981;
43. M. V. Klein, ***Optics***, John Wiley & Sons, Inc., New York, 1970;
44. M. Kline, I. W. Kay, ***Electromagnetic Theory and Geometrical Optics***, John Willey & Sons, Inc., New York, 1965;
45. Yu. A. Kravțov, Yu. I. Orlov, ***Geometriceskaia optika neodnorodnîh sred***, Izd. Nauka, Moskva, 1980;
46. E. E. Kriezis și alții, ***Electromagnetics and Optics,*** World Scientific, Singapore, 1992;
47. M. Lanchenaud, ***Instruments d'Optique***, Dunod, Paris, 1976;
48. L. Landau, E. Lifchitz, ***Théorie du champ,*** Chap. VII, ***Propagation de la lumière***, Editions Mir, Moscou, 1966;
49. G. S. Landsberg, ***Optica***, prima parte, Editura Tehnică, București, 1958;
50. L. Levi, ***Applied Optics, A Guide to Optical Design***, John Wiley & Sons, Inc., New York, 1968;
51. R. S. Longhurst, ***Geometrical and Physical Optics***, Longmans, Green and Co., Ltd., London, 1984;
52. D. J. Lowell, ***Optical anecdotes***, SPIE, Washington, 1984;
53. H. Lumbroso, ***Optique géométrique et ondulatoire*** (98 probèmes resolus), Dunod, Paris, 1996;
54. R. K. Luneburg, ***Mathematical Theory of Optics***, University of California Press, Berkeley, 1964;
55. E. Mach, ***The Principles of Physical Optics. An Historical and Philosophical Treatment***, Dover Publication, Inc., New York, 1953;
56. D. Marcuse, ***Light Transmission Optics***, Van Nostrand Reinhold, New York, 1972;
57. A. Maréchal, ***Optique géométrique générale***, in ***Handbuch der Physik***, herdansgegebem von S. Flügge, Band XXIV, ***Grundlagen der Optik***, Springer - Verlag, Berlin, 1956;
58. L. C. Martin, ***Geometrical Optics***, Philosophical Library, Inc., New York, 1956;
59. L. C. Martin, W. T. Welford, ***Technical Optics***, Pitman & Sons, Ltd., London, 1966;
60. A. N. Matveev, ***Optics,*** Mir Publishers, Moskow, 1988;
61. M. May, A-M. Cazabat, ***Optique - Cours et problèmes résolues***, Dunod, Paris, 1996;

62. M. May, ***Introduction à l'Optique***, Dunod, Paris, 1996;
63. J. R. Meyer - Arendt, ***Introduction to Classical and Modern Optics***, Prentice - Hall, Inc., London, 1972;
64. G. C. Moisil, E. Curatu, ***Optică, teorie și aplicații***, Editura Tehnică, București, 1986;
65. K. D. Möller, ***Optics***, University Science Books, Mill Valley, 1988;
66. G. Nemeș, I. Teodorescu, M. Nemeș, ***Optica în spațiul fazelor, teorie și aplicații***, Editura Academiei Române, București, (va apare);
67. Isaac Newton, ***Optica***, Editura Academiei Române, București, 1970, în traducerea Prof. Victor Marian;
68. I. Nicoară, ***Optică***, Tipografia Universității din Timișoara, 1988;
69. A. Nussbaum, ***Geometric Optics***, Addison - Wesley Publ. Co., Inc., Reading, Mass., 1968;
70. G. H. Owyang, ***Foundations of Optical Waveguides***, Elsevier, New York, 1981;
71. F. L. Pedrotti, L. S. Pedrotti, ***Introduction to Optics***, Prentice-Hall, 1993;
72. J-P. Pérez, ***Optique*** (***Fondements et applications***), Masson, Paris, 1996;
73. V. Pop, ***Optica***, Vol. II, Universitatea "Al. I. Cuza", Iași, partea I, 1983, partea a doua, 1986;
74. I. I. Popescu, ***Optica, 1. Optica geometrică***, Tipografia Univ. din București, 1988;
75. I. I. Popescu, E. I. Toader, ***Optica***, Ed. Științifică și Enciclopedică, București, 1989;
76. I. M. Popescu, ***Teoria electromagnetică macroscopică a luminii***, Editura Științifică și Enciclopedică, București, 1986;
77. J. P. Provost, P. Provost, ***Optique, vol.1, Optique et Principe de Fermat, vol.3, Exercices et problèmes d'optique géometrique***, Cedic/Fernand Nathan, Paris, 1980;
78. V. Ronchi, ***The Nature of Light, an Historical Survey***, Harvard University Press, 1970;
79. B. Rossi, ***Optics***, Addison-Wesley Publishing Company, Reading, 1962;
80. M. P. Silverman, ***Waves and grains (Reflections on Light and Learning)***, Princeton University Press, Princenton, 1998;
81. D. Sivoukhine, ***Cours de Physique Générale***, tome IV, ***Optique***, première partie, Editions Mir, Moscou, 1984;
82. F. G. Smith, J. H. Thomson, ***Optics***, Wiley, New York, 1988;
83. F. G. Smith, T. A. King, ***Optics and Photonics (An Introduction)***, Wiley&Sons,Chichester, 2000;
84. M. I. Sobel, ***Light***, The Univ. of Chicago Press, 1987;
85. A. Sommerfeld, ***Lectures on Theoretical Physics***, Vol. IV, ***Optics***, Academic Press Inc., New York, 1954;
86. O. N. Stavroudis, ***The Optics of Rays, Wavefronts and Caustics***, Academic Press, New York and London, 1972;
87. P. Ștețiu, ***Optica I, Optica Geometrică***, Tipografia Universității din Cluj-Napoca, 1987;
88. J. L. Synge, ***Geometrical Optics, An Introduction to Hamilton's Method***, Cambridge University Press, 1962;
89. E. I. Toader, ***Aparate optice***, Ed. Științifică, București, 1995;
90. R. Țițeica, I. I. Popescu, ***Fizica Generală***, Vol. II, Editura Tehnică, București, 1973;
91. F. Uliu, ***Optică și Spectroscopie***, partea I, ***Optica Geometrică***, Tipografia Universității din Craiova, 1973; ***Culegere de probleme de optică***, Tipografia Universității din Craiova, 1979;
92. F. Uliu, ***Curcubeul - De la mit la adevăr***, Editura SITECH, Craiova, 1994;

93. F. Uliu, ***Istoria curcubeului - de la Noe la Mie***, Editurile Emia (Deva) și Universitaria (Craiova), 2005;
94. F. Uliu, ***Probleme alese de fizică, vol.II*** (***Optică și relativitate***), Ed. Radical, Craiova, 1997;
95. H. G. Unger, ***Planar Optical Waveguides and Fibres***, Oxford Univ. Press, 1980;
96. I. Ursu, ***Optica***, Litografia Universității din Cluj, Cluj, 1953;
97. M. Young, ***Optics and Lasers***, Fifth Edition, Springer, Berlin, 2000;
98. W. T. Welford, ***Geometrical Optics***, North Holland Publ., Co., Amsterdam, 1962; ***Aberations of the Symetrical Optical Systems***, Academic Press, London, 1974;
99. H. G. Zimmer, ***Geometrical Optics***, Springer - Verlag, Berlin, 1970;
100. * * * ***Sbornic zadaci po obșcemu kursu fiziki (Optica)***, pod redacția D. V. Sivoukhina, Izd. Nauka, Moskva, 1977;
101. * * * ***Melles Griot, Optics Guide***, 5, 1990.





Contemporary
Literature Press
http://editura.mttlc.ro
The University of Bucharest. 2017